## REMARKS

Applicants note that Claims 245, 248-251, 253-255, 260, 264, 268, 270, 270, 272, 284, 288-290, 296, 299, 303 304, 308-313 and 318-326 are pending in the above-referenced application. Claims 318-323 have been withdrawn. As will be discussed in further detail below, claims 265, 296, 299, 308 and 325 have been amended to more distinctly claim that which Applicants regard as the invention. These amendments have been made to advance prosecution. Applicants do reserve the right to file subsequent continuation and/or divisional applications on canceled subject matter.

Specifically, claim 265 has been amended to recite that the snRNA is U1, U2 and U4 snRNA. The recitation of U1 and U2 is expressly supported by the specification in the paragraph bridging pages 103 and 104 and the line 9 of page 104. Applicants note that there is also implicit support on page 104, line 9 since it states "This invention should be applicable to other species of snRNA including U2." Furthermore as discussed during the interview on June 24, 2009, the following is stated on page 9, lines 4-13:

> U1, U2 and other snRNAs are nuclear-localized RNA molecules complexed with protein molecules. (Dahlberg and Lund 1988 in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, M. Birnstiel, Ed., Springer Verlag, Heidelberg, p38; Zieve and Sautereau 1990 Biochemistry and Molecular Biology 25; 1, all of which are incorporated herein by reference).
>
> The various promoters for U1, U2 and other snRNA operons are very strong and produce large amounts of RNA. U1 and other snRNAs have signals for export to the cytoplasm where specific proteins are complexed before reimportation to the nucleus as snRNPs.

The reference Zieve and Sautereau which as stated in the specification is incorporated by reference shows in Figure 1 U1, U2 and U4. A copy of Figure 1 of Zieve and Suatereau is attached hereto as Appendix A.

Claim 296 has been amended to recite that the construct comprises U1 or U2 snRNA or both. This amendment was made to insert "snRNA" after U2. No new matter was added and the amendment is supported by the specification.

Applicants further note that claims 299, 308 and 325 have been amended to recite that the specific nucleic acid produced is complementary with a specific portion of one or more viral RNAs or binds to a specific viral protein. This amendment merely deletes the phrase "or cellular". Thus, no new matter was added and these amendments are supported by the specification. Further, in view of the amendments of claims 299, 308 and 325, claims 309-311 and 324 have been canceled.

## I. SUBSTANCE OF INTERVIEW

First, Applicants would like to thank Examiner J. Zara for her time and thoughtful suggestions during the interview with Applicants representative, Cheryl H. Agris and one of the inventors, Dr. James Donegan on June 24, 2009. The Substance of the Interview is set forth below.

### A. Brief Description of any Exhibit Shown or any Demonstration Conducted

Applicants submitted Figures 37-43 and 46 of the specification since these pages and figures were referred to during the interview.

### B. Identification of Claims Discussed

Claims 245, 265, 299, 325 and 326 were discussed.

### C. Identification of Specific Prior Art Discussed

As will be set forth in further detail below, Calabretta et al. was discussed with respect to the rejections under 35 USC §103.

### D. Identification of Principal Proposed Amendments of a Substantive Nature Discussed

Possible amendments to claims 265, 299 and 325 were discussed.

**E. Identification of General Thrust of Principal Arguments presented to the examiner**

An adequate description has been provided to support the pending claims. No new matter is contained in the pending claims. Further, none of the pending claims are obvious over the cited references.

**F. A General Indication of Any other Pertinent Matters Discussed**

Possible support for amended claim 265 reciting U1, U2 or U4 snRNAs was discussed.

**G. General Results or Outcome of the Interview**

Applicants agreed to submit arguments to support assertions of adequate written description and further agreed to point out with specificity sections of the specification that support the instant claims. Furthermore, Applicants will set forth arguments as to why the claims are not obvious over the cited references.

**II. The Rejection Under 35 USC §112, First Paragraph**

Claims 245, 248-251, 253-255, 260, 264, 265, 268, 270, 272, 284, 288-290, 296, 299, 303, 304, 308-313, 325 and 326 are rejected under 35 U.S.C. §112, first paragraph for lack of written description. The Office Action specific states:

> The claims are drawn to an isolated multi-cassette nucleic acid construct comprising at least three promoters, and which optionally comprises a nuclear localization sequence comprising a portion of snRNA comprising sequences for at least two stem loops present at the 3' end of native snRNA, a reimportation, and an antisense nucleic acid sequence replacing stem-loop formation of native snRNA, and which nucleic acid construct produces, upon introduction into any eukaryotic cell, at least one specific nucleic acid from each promoter or initiator, which upon insertion into a eukaryotic cell produces more than one specific nucleic acid, each such specific nucleic acid so produced being substantially non-homologous with each other and being either complementary with a specific portion of one or more viral or cellular RNAs in a cell or binds to a specific viral or cellular protein, which virus is optionally HIV, wherein each specific

nucleic acid binds to a different target nucleic acid sequence, and the specific nucleic acid binds to a specific cellular protein comprising a localizing protein or a decoy protein.

The specification and claims do not adequately describe the various genera comprising i.) any snRNA comprising sequences for stem loops present at the 3' end of any native snRNA, and which comprise any reimportation signal or which comprise any antisense replacing sequences that participated in stem-loop formation in the native form of any snRNA; ii.) any cellular protein comprising any nuclear localizing protein or cytoplasmic localizing protein; iii.) any decoy protein binding to any protein required for viral assembly or viral replication.

The instant disclosure, at the time of filing, does not provide adequate number of species for the broad genera claimed. The specification teaches the human U1 operon, and elimination of 49 base sequence involved in the formation of A and B loops formed by U1. The specification also teaches a three segment, triple operon constructs comprising either three U1 promoters or three T7 promoters, and antisense targeting HIV 5' common leader, the TAT/REV coding sequence and the splice acceptor site for TAT/REV of HIV.

The disclosure of these constructs, however, is insufficient to teach or adequately describe a representative number of species for the broad genera of nucleic acids constructs claimed, such that the common attributes or characteristics concisely identifying members of each proposed genus are exemplified, and further whereby any primary nucleic acid construct comprising any primary nucleic acid sequence is introduced into any eukaryctic cells and acts as a template for the synthesis of any secondary nucleic acid for the synthesis of any gene product, which nucleic acid construct comprises any snRNA comprising sequences for stem loops present at the 3' end of any native snRNA, and which comprise i.) any reimportation signal or which comprise any antisense replacing sequences that participated in stem-loop formation in the native form of any snRNA; ii.) any cellular protein comprising any nuclear localizing protein or cytoplasmic localizing protein; iii.) any decoy protein binding to any protein required for viral assembly or viral replication. The general knowledge and level of skill in the art at the time of filing do not supplement the omitted description because specific, not general, guidance is what is needed to provide a representative number of species for the broad array of nucleic acid constructs claimed.

> Since the disclosure and the prior art, at the time of filing, fail to describe the common attributes or characteristics concisely identifying members of the proposed genera of compounds claimed, or fail to provide and adequate number of species for the broad genera claimed, the description provided for this very broad genera of compounds is insufficient. One of skill in the art would reasonably conclude that the disclosure fails to provide a representative number of species to describe the very broad genera claimed.

Applicants respectfully traverse the rejection. It is Applicants view that there is adequate support for the subject matter recited in the pending claims. There are three sets of claims, (1) claims 245, 246-255, 260 and 264, (2) claims 265, 268, 270, 272, 284, 288-290 and 296 and (3) claims 299, 303, 304, 308-313, and 324-326. Each will be addressed below.

**A. Claims 245, 246-255, 260 and 264**

Below is a table summarizing the support for claim 245:

| Recitation | Support |
|---|---|
| A composition comprising (a) an isolated primary nucleic acid construct comprising a primary nucleic acid, which upon introduction into a eukaryotic cell acts as a template for the synthesis of a secondary nucleic acid which acts as a template for the synthesis of a gene product, selected from the group consisting of a sense and antisense nucleic acid in said eukaryotic cell, wherein said secondary nucleic acid or said gene product **does not act as a template for the synthesis of said primary nucleic acid** and | p.92, par. 1, p. 93, par. 3, Examples 23-25, diagrammatically depicted in Figs. 37-39 |
| b) a signal processing sequence. | Page 96, lines 12-15: "When the above-described compositions further comprise a signal processing sequence, such sequences can be selected from a promoter, an initiator... " |

Furthermore, as asserted in the response to the previous Office Action dated April 16, 2008, the specification on page 92, paragraph 1 states:

> The generation or formation of a Production Center from a Primary Nucleic Acid Construct or the generation or formation of a Production Center from another Production Center. However, production centers cannot produce a Primary Nucleic Acid Construct.

A production center is defined on page 91, lines 14-19 as follows:

> As used herein, the term production center is intended to cover secondary nucleic acid components which can be produced from a primary nucleic acid construct. Also covered are a tertiary nucleic acid which could be produced from the secondary nucleic acid component, as well as any nucleic acid product which may be produced from the secondary nucleic acid component.

Furthermore, it is stated on page 93, lines 11-17.

> Thus, a significant embodiment of this invention concerns a composition comprising a primary nucleic acid component which upon introduction into a cell produces a secondary nucleic acid component which is capable of producing a nucleic acid product, or a tertiary nucleic acid component, or both. The secondary and tertiary nucleic components and the nucleic acid product are incapable of producing the primary nucleic acid component.

The term "nucleic acid component" is defined in the paragraph bridging pages 93 and 94 as:

> In the present composition, the primary nucleic acid component can comprise a nucleic acid, a nucleic acid construct, a nucleic acid conjugate, a virus, a viral fragment, a viral vector, a viroid, a phage, a phage vector, a plasmid, a plasmid vector, a bacterium and a bacterial fragment or combinations of any of these.

Thus, the term "nucleic acid component" encompasses a nucleic acid construct or nucleic acid.

Specific examples of production centers are provided in Figures 36-40. A description of Figures 36-40 is provided in Examples 22 (describes Figure 36), 23 (describes Figures 37 and 38), 24 (describes Figure 39) and 25 (describes Figure 40).

Claims 246-255, 260 and 264 depend from claim 245. Therefore, arguments made with respect to claims 245 would apply to these claims as well.

**B. Claims 265, 268, 270, 272, 284, 288-290 and 296**

Applicants further assert that there is adequate written description of claims 265, 268, 270, 272, 284, 288-290 and 296. First, in response to the assertion that the specification and claims do not adequately describe the various genera comprising any snRNA comprising sequences for stem loops present at the 3' end of any native snRNA. Applicants note that there are only a limited number of snRNAs (with the major species classically considered to be U1-U7). At the time of the filing, consensus secondary structures have been generated for U1, U2, U4, U5 and U6. Secondly, only some of them have a reimportation signal: U1, U2, U4 and U5. Further, as noted earlier in the response, in order to advance prosecution, claim 265 has been amended to recite that the snRNA is U1, U2 or U4 snRNA. Although, only U1 was used, there would be a minimal amount of effort required to apply the present methods to U2 or U4. The insertion of appropriate sequences could be done by a variety of method besides the methods actually employed. For instance, a set of oligonucleotides could be synthesized that could be hybridized together and ligated to create an snRNA cassette with a restriction site or sites that would be convenient for inserting a sequence of interest. Conversely, rather that a universal cassette, a fusion product could be custom-made, where a set of oligonucleotides is synthesized, hybridized and ligated together that includes the anti-sense sequence of interest.

With regard to "any anti-sense RNA", Applicants have basically described the use of the snRNA as a scaffold where the anti-sense sequence of choice may be inserted. This would preserve the desirable properties that are the physical property of snRNAs and allow these properties to be shared by anti-sense transcripts. As such the invention would be independent of a particular anti-sense RNA transcript and the method could be used for any anti-sequence that was of interest to the user. This is the

same way that a novel expression vector should be patentable regardless of the particular sequence coding sequence that is inserted for expression.

Claims 268, 270, 272, 284, 288-290 and 296 depend from claim 265. Therefore, arguments made with respect to claim 265 would apply to these claims as well.

**C. Claims 299, 303, 304, 308-313, and 324-326**

Applicants with respect to claims 299, 308 and 325 notes that these claims have been amended to recite that the specific nucleic acids produced are complementary with a specific portion of one or more **viral** RNAs in a cell or binds to a specific **viral** protein. There is adequate support for this recitation. Further, Applicants note that in view of the amendment of claim 299, claims 309-311 and 324 have been canceled. Additionally, claim 326 specifically recites that that nucleic acid produced is complementary with a specific portion of one or more HIV RNAs in a cell or binds to a specific HIV.

With respect to assertions made regarding snRNAs, as noted above, there are only a limited number of snRNA's (with the major species classically considered to be U1-U7). At the time of the filing, consensus secondary structures have been generated for U1, U2, U4, U5 and U6. Further, the insertion of appropriate sequences could be accomplished by a variety of methods besides the methods actually employed (see, for example, Figure 46).

Claims 303, 304, 312-313 depend from claim 299. Thus arguments made with respect to claim 299 are applicable to these claims as well.

In view of the above arguments and amendments of claim 265, 296, 299, 208 and 325, the rejection of the pending claims over 35 USC §112, first paragraph have been overcome. Therefore, Applicants respectfully request that these rejections be withdrawn.

**III. The Rejections Under 35 USC 6103**

Claims 245, 248-251, 253-255, 260, 264, 299, 303, 304, 308-313, 325, 326 have been rejected under 35 U.S.C. §103(a) as being unpatentable over Calabretta et al. (USPN 5,734,039), in view of Binkley et al. (Nucleic Acids Research, 1995, Vol. 23, No.

16, pages 31 98-3205), the combination further in view of Craig et al. (WO 95108635) and Alul et al (USPN 5,532,130). The Office Action specifically states:

> It would have been obvious to incorporate RNA oligonucleotides that bind to proteins, as taught by Binkley et al. in place of the antisense oligonucleotides taught in the system of Calabretta et al. One would have been motivated to incorporate RNA oligonucleotides that bind to proteins instead of the antisense oligonucleotides in the system of Calabretta et al. because Binkley et al. teach that high affinity RNA ligands to proteins, such as NGF that localizes NGF-sensitive growing axons, can be easily isolated using the SELEX procedure and teach that such RNAs may furnish useful diagnostic tools for the study of proteins. Since both types of nucleic acid oligonucleotides are used to determine binding interactions, as evidenced by the teachings of Calabretta et al. and Binkley et al., one would have been motivated to express the RNA ligands taught by Binkley et al. in the system of Calabretta et al.
>
> One would have a reasonable expectation of success given that each of the nucleic acid molecules were known to bind with target molecules in a sequence specific manner, as evidenced by Calabretta et al. and Binkley et al. One would have a reasonable expectation of success to express the protein binding RNA molecules of Binkley et al. in the dual system of Calabretta et al., with the advantage of producing two different binding molecules at once.
>
> It also would have been obvious to use the SELEX method to assay for RNA molecules that it would have been obvious to incorporate RNA oligonucleotides that bind to proteins, as taught by Binkley et al. in place of the antisense oligonucleotides taught in the system of Calabretta et al. One of ordinary skill would have been motivated to design and synthesize antisense that target and inhibit HIV proteins because in the search for potential therapeutics to inhibit HIV infections, as taught previously by many in the art, including Alul et al. One would have been motivated to screen for resultant RNA aptamers against a decoy protein because Binkley et al. teach that high affinity RNA ligands to proteins can be easily isolated using the SELEX procedure and teach that such RNAs may furnish useful diagnostic tools for the study of proteins. Since Craig et al. teach that decoy proteins are proteins that are useful to serve as a mutant that is capable of binding to a preferred site but yet is incapable of activating transcription, one would have been motivated to

> use the SELEX method of Binkley et al. to identify RNA ligands to any known protein, such as the decoy proteins of Craig et al.
>
> One would have a reasonable expectation of success given that Craig et al. teach the benefits of decoy proteins and Binkley et al. teach assaying for RNA aptamers to proteins and teach a method (SELEX) that is widely use to identify RNA molecules that bind to known proteins……
>
> …it would have been obvious to construct a nucleic acid construct with more than two different promoters driving expression of different gene products. This would have involved nothing more than routine experimentation at the time of the instant invention, and relying on the prior teachings of Calabretta and Akil et al for the utilizing these constructs to target HIV RNA.

Applicants respectfully traverse the rejection. First, Applicants assert that this rejection should not even be applicable with respect to claims 245, 248-251, 253-255, 260, 264, given that these claims are not focused on multi-promoter cassettes. Furthermore, the Calabretta reference in combination with other secondary references has in previous Office Actions been applied to claims 299, 303, 304, 308-313, 325, 326.

Applicants again traverse the rejection with respect to claims 299, 303, 304, 308-313, 325 and 326. With regard to Calabretta, Applicants wish to reiterate the position that their teachings were based on the idea that two different cellular locales (cytoplasmic and nuclear) demanded two different promoters. Regardless of the nature of the particular sequences being expressed from each promoter of Calabretta (since this is the only element that is being added by the other references), Applicants again disagree that any particular motivation exists for Calabretta to expand to more than two promoters since both of the essential two locales are already covered by the use of two promoters. It isn't just that Calabretta uses "two different genes to a cell" but rather that they go to great length to explain that the cytoplasmic as well as the nuclear locales should be separately targeted. There is no third locale other than these two portions of the cell, thereby offering no particular incentive to add a third promoter. Even if it was desired to target more than one cytoplasmic target, the art of the time of Calabretta only describes the use of a single multivalent anti-sense RNA rather than separate

transcripts. Applicants note that claims 299 and dependent claims 303, 304, 308 and 312-313 and claim 326 recite that the construct comprises at least three promoters.

Claim 325 recites that the promoter is either an snRNA promoter or bacteriophage promoter. Calabretta et al. does not teach either snRNA promoters or bacteriophage promoters. No direction was given in Calbretta et al. that would motivate one of ordinary skill in the art to choose these promoters.

The other cited references would be of limited significance. Binkley et al. merely teaches molecules that may bind to cellular protein. Craig et al. merely teaches expression of a viral decoy protein. In Applicants view, it would not be obvious to combine all of these references. As noted above, combining Binkley et al. with Calabretta et al. would at best provide a construct that expresses two specific RNA sequences that binds to a cellular protein. Given that claims 299, 308 and 325 have been amended to recite that the specific nucleic acid binds to viral protein and given that claim 326 is directed to an HIV RNA that binds to HIV protein, Binkley should not even be applied. Further, there was no suggestion regarding combining Craig et al. with the other two. Craig et al. merely teaches the cloning of a protein and its therapeutic uses. There is no teaching regarding binding to a specific nucleic acid or facilitate transport.

Alul merely discloses anti-sense sequences containing 2'-5' linkages and their uses as therapeutics. The Background of the invention merely summarizes the state of the art with respect to the effect of various oligonucleotide analogs on gene expressions. However, in Applicants' view, it would not have been obvious to combine Alul with the other cited references since the 2'-5' oligonucleotides are not incorporated into constructs but are used as stand alone therapeutics. The claims recite a construct that would be used as a template for synthesis when present in a cell. It's difficult to imagine circumstances where incorporation of the 2'-5' nucleotides of Alul would take place within a cell.

In view of the above arguments, Applicants assert that the rejections under 35 USC §103 have been overcome. Therefore, Applicants respectfully request that the rejections be withdrawn.

## SUMMARY AND CONCLUSIONS

It is Applicants belief that the pending claims are in condition for allowance. However, if a telephone conversation would further the prosecution of the present application, Applicants' undersigned attorney request that he be contacted at the number provided below.

Respectfully submitted,

/Cheryl H Agris/

Dated: July 7, 2009

Cheryl H. Agris, Reg. No. 34,086
Telephone No. (914) 712-0093

# APPENDIX A

# Cell Biology of the snRNP Particles

*Gary W. Zieve and Roger A. Sauterer*

Referee: Robert L. Margolis
Department of Basic Sciences
Hutchinson Cancer Research
Seattle, Washington

## I. STRUCTURE, COMPOSITION, and GENETICS of snRNPs

### A. General Characteristics of snRNPs

The small nuclear RNAs (snRNAs) are a ubiquitous class of low-molecular-weight RNAs found in ribonucleoprotein particles, the snRNP particles, in the cell nucleus. The snRNAs are stable and with limited exceptions have the diagnostic hypermethylated, 2,2,7 trimethylguanosine 5′ cap and assemble with a common set of six snRNP core proteins in addition to one or more snRNP-specific proteins. The snRNP particles function in RNA processing in the nucleus, including pre-mRNA splicing and pre-mRNA 3′ end processing in the nucleoplasm. The snRNPs assemble into complexes with their processing substrates where several of the snRNAs base pair with conserved sequence motifs in the substrates and in other snRNAs.[1-5] In mammalian cells there are six major snRNAs, named U1 through U6, ranging in size from 106 to 217 nucleotides[6] (see Table I and Figures I and 2), which are present in approximately $1 \times 10^6$ to $1 \times 10^6$ copies per cell nucleus in human cells. A growing family of less-abundant snRNAs (U7 to U12), identified by their trimethylguanosine 5′ caps and their immunoprecipitation by the anti-Sm class of autoimmune antibodies from patients with systemic lupus erythematosus (SLE) (discussed in detail later), are being described[6,8,9] (see Table 1). Two dozen snRNAs with diagnostic trimethylguanosine 5′ caps have been described in yeast cells and six are clearly homologous to the mammalian U1 - U6 snRNAs in structure and function[5,10,11] (see Tables 2 and 3). The abundance of the snRNAs in yeast is at least tenfold lower than that in mammalian cells, which correlates with the substantially fewer transcribed introns and subsequent RNA processing in the yeast nucleus.[10] The snRNAs are transcribed by RNA polymerase II, except for U6, which is transcribed by RNA polymerase III.[12-14]

The snRNAs exist in the cell in the form of ribonucleoprotein complexes, the snRNPs, which sediment at 10S to 14S and are visible in negatively stained electron micrographs as irregular spheroids approximately 10 nm across.[15,16] With the exception of U6 and the nucleolar snRNP U3, the snRNPs in mammalian cells share a common core of six polypeptides (B 28 kDa, B′ 18 kDa, D 16 kDa, E 13 kDa, F 12 kDa, and G 11 kDa), which recognizes a conserved sequence motif, the Sm binding site, and is assembled into a particle of $B_1B'D_1D_2EFG$[2,17] (see Table 4 and Figures 1 and 2). In addition, each snRNA is associated with one to three specific polypeptides to form particles that are approximately 80% protein. Although the U6 snRNA is not assembled with the common core of snRNP proteins, it is precipitated by antibodies against the core protein because it is base-paired with U4 in a single particle.[18,19] There are both species and tissue differences in the B protein, including a neural-specific form of the B protein, the N protein, and in human, but not rodent cells, the B′ protein, a variant of the B protein.[17,20]

Autoimmune antibodies have provided powerful tools for studying the snRNP particles (for summary, see Table 5). The Sm and (U1)RNP antisera found in over 40% of the patients with SLE recognize protein components of the snRNP particles.[21] Antigenic determinants on the 28-kDa B and 16-kDa D proteins of the snRNP core are the major epitopes recognized by the anti-Sm class of autoantibodies, and the anti-(U1)RNP autoantibodies recognize determinants on the U1-specific proteins[22] (Figure 3). A growing family of rare autoimmune specificities have been identified that recognize snRNP-specific proteins on other snRNP particles, including antibodies to fillagrin, a 34-kDa protein nucleolar protein that is a component of the U3 snRNP particle.[23,24]

SnRNP particles assemble in the cytoplasm, where newly transcribed snRNAs associate with snRNP core proteins present in large pools of partially assembled RNA-free intermediates, before returning permanently to the interphase nucleus.[25] In the cytoplasm, the D, E, F, and G snRNP core proteins preassemble into a 6S RNA-free complex of $D_2EFG$, and the B protein assembles into a heterogeneous set of RNA-free homo-oligomers (see Figure 6). Newly transcribed snRNAs assemble with the 6S $D_2EFG$ particle followed by two copies of the D′ protein and then two copies of the B protein and the snRNP-specific proteins. Cytoplasmic maturation of the snRNAs includes hypermethylation of the monomethylated cap to a trimethylated cap and cleavage of extra 3′ terminal nucleotides.[26-30]

In the cell nucleus, the mature snRNP particles function in the processing of newly transcribed RNA, where several of the snRNAs base pair with conserved sequence motifs in their processing substrates, and possibly with other snRNAs during these processing events. The snRNPs participate in both the splicing and 3′ end processing of pre-mRNA, and the processing of pre-rRNA. In the nucleoplasm, the U1, U2, U4/U6, and U5 snRNPs assemble into the spliceosome (Figure 5) that removes introns from pre-mRNA by splicing. U1 and U2 base pair with

G. W. Zieve received a B.S. degree from the California Institute of Technology in Pasadena, California, and a Ph.D. degree from the Massachusetts Institute of Technology in Cambridge. Dr. Zieve is currently an Associate Professor in the Department of Anatomical Sciences at the State University of New York at Stony Brook. R. Sauterer received a B.A. degree from Oberlin College in Oberlin, Ohio, and a Ph.D. degree from the State University of New York at Stony Brook. Dr. Sauterer is currently a Research Associate in the Department of Anatomy at the Albert Einstein College of Medicine in New York.

Table 1
The SnRNAs in Mammalian Cells[1,2,4-8]

| | Size Nucleotides | Abundance × 10⁶ | 5′ End |
|---|---|---|---|
| U1 | 165 | 1.0 | m3GpppAmUmA |
| U2 | 189 | 0.9 | m3GpppAmUmC |
| U3 | 216 | 0.2 | m3GpppAmAmC |
| U4 | 139 | 0.2 | m3GpppAmGmC |
| U5 | 117 | 0.2 | m3GpppAmUmA |
| U6 | 107 | 0.4 | $CH_3$pppGUG |
| U7 | 56 | <0.05 | m3GpppAUC (sea urchin) |
| U8 | 139 | <0.05 | m3GpppAmUmC |
| U9 | 130 | <0.05 | m3GpppA |
| U10 | 60 | <0.05 | m3GpppA |
| U11 | 131 | <0.05 | m3GpppA |
| U12 | 150 | <0.05 | m3GpppN |

conserved sequence motifs at the 5′ splice site and the branch point, respectively, and the U1 snRNP recognizes both the 5′ and 3′ splice sites and may have a fundamental role in aligning the two sites for the splicing reaction. The U4/U6 and U5 snRNPs preassemble into a single 25S particle that binds to U1, U2, and the substrate to form the mature spliceosome. The intron is then removed by two concerted cleavage and ligation reactions.[4,9,11,12] The U7 snRNP base pairs with a conserved sequence in the 3′ end of histone pre-mRNA and is an essential cofactor for the cleavage that creates the mature nonadenylated 3′ end. The U11 snRNP is present in the polyadenylation complex that processes the 3′ end of most pre-mRNAs; however, there is no evidence of base pairing with the substrate.[13-16] Species U3 and U8 are localized in the nucleolus, where U3 base pairs with the 28S rRNA precursor in a still unidentified region and functions in pre-rRNA processing.[18,23-26] This review covers (1) the structure of the snRNAs and their genes, (2) the structure of the snRNP particles and the regulation of snRNP assembly, and (3) the function of mature snRNP particles in the nucleus.

## B. Evolutionary Conservation of snRNAs

In vertebrates, the snRNAs are highly conserved, with sequence differences of less than 10% between birds and mammals.[1,4] U1 and U2 are the most abundant snRNAs, with U3, U4, U5, and U6 present in approximately a tenfold lower abundance. The growing family of other snRNAs, U7 to U12, are another tenfold lower in abundance (Table 1). Even where primary sequences diverge between widely separated organisms, the secondary structures of snRNAs are strongly conserved in species as different as yeast and humans.[4] The snRNAs have several conserved sequence motifs, including a 6- to 12-nucleotide, single-stranded sequence, $PuA(U)_nGPu$, $n = 2-6$, found in almost all the snRNAs, that directs the binding of the common set of snRNP core proteins reactive with the Sm antisera and is called the Sm binding site[28,29] (see Figure 1). There are also sequence domains in several snRNAs capable of base pairing with their processing substrates and other snRNAs.

The major snRNAs from *Drosophila* cells have similar mobilities to the mammalian U1 to U6 snRNAs.[30-40] With the exception of U3, these snRNAs are immunoprecipitable by anti-Sm antisera, which recognizes the common core of mammalian snRNP proteins, and sequence analysis shows a 75% homology with mammalian counterparts. The secondary structure of the various *Drosophila* snRNAs is in general highly conserved compared with the mammalian snRNAs[31] (see Figure 1C). Sequence homology is particularly strong in specific regions that probably are necessary to their function, such as the 5′ end of U1, the putative Sm antigen binding site, and regions in the stem and loop of the different snRNAs.[31,40] SnRNAs showing homology to the major mammalian snRNAs have been found in microorganisms such as *Amoeba*, *Tetrahymena*, *Bombyx*, and *Dictyostelium*, as well as in fungi and higher plants.[3] In *C. elegans*, the common leader that is found on a large set of mRNAs is initially transcribed and processed like a snRNA. During the initial splicing event in the nucleus, this snRNA serves as the first exon and its 5′ 22 nucleotides spliced to the substrate to form the common spliced leader.[41-43]

### 1. Low-Abundance Mammalian snRNAs (U7 to U12)

Recently, several minor snRNAs have been identified in mammalian cells because of their diagnostic trimethylguanosine cap (Table 1). These RNAs are much less abundant than the U1 to U6 snRNAs.[7-9]

The U7 snRNA is a 57-nucleotide snRNA that is 0.2% of the abundance of the U1 snRNA. This snRNA is an essential cofactor for the 3′ end processing of histone pre-mRNA. A 13-nucleotide-long sequence (from 13 to 26 nucleotides from the 5′ end) base pairs with a universally conserved block at the 3′ end of histone pre-mRNA and helps specify the site of 3′ end cleavage.[44-46] U7 is immunoprecipitable by Sm antisera; however, its Sm binding site is divergent from the conserved Sm binding consensus sequence found in other snRNAs[44-46] (see Figure 1F).

Five other small, low-abundance RNAs (U8 to U12), all immunoprecipitable by anti-Sm and anti-trimethylguanosine cap antibodies, are characterized in rat and human cells.[7-9] U8, U11, and U12 are sequenced and each has a trimethylguanosine cap and a Sm binding site, although the U8 Sm sequence diverges slightly from the consensus. U8 snRNA is present only in 25,000 copies per cell, and shows high (95%) sequence conservation between rat and human cells.[7] It is localized predominantly in the nucleolus, as is U3. Both U5 and U10 show unique fingerprints unlike other known snRNAs.[9] U11 and U12 are sequenced and contain the typical Sm binding site and possess a trimethylguanosine 5′ cap.[8] The U11 snRNP cofractionates with an essential activity for *in vitro* polyadenylation of pre-mRNA; however, there is no identified sequence complementarity with conserved sequence motifs in the pre-mRNA substrate.[16] A Sm precipitable complex binds to the conserved polyadenylation sequence motif and protects it from nuclease digestion *in vitro*, but it is not known if it is a snRNP or a RNA-free particle.[15] The nucleolar location

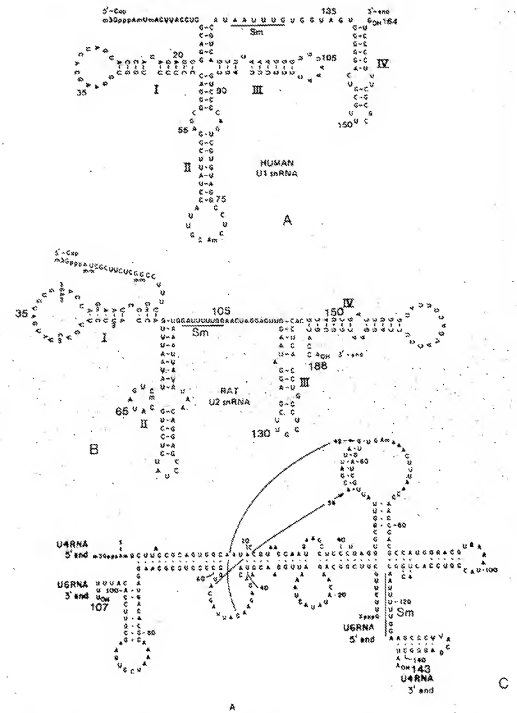


FIGURE 1. Sequences and predicted secondary structures of the major snRNAs. Sequences and predicted secondary structures are illustrated for human U1 (A), rat U2 (B), human U3(C), *Drosophila* U4/U6 (D), human U5(E), and sea urchin U7(F) snRNAs. Stem loops are numbered and the Sm binding site is underlined. Methylated nucleotides are indicated with a m and pseudouridine as a ψ. Model for the U4/U6 is the extensively base-paired model of Hashimoto and Steitz.[34]

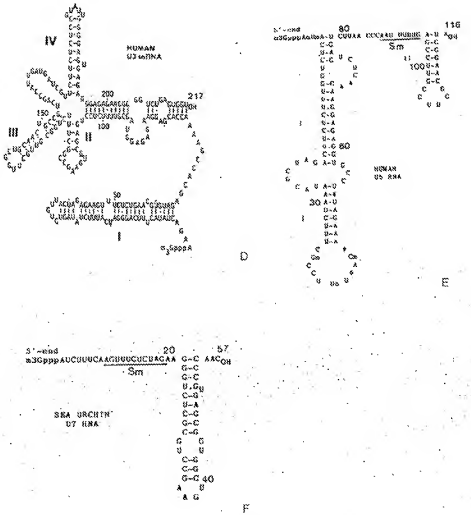


FIGURE 1B

of U8 suggests that it may play a role in ribosomal RNA processing. The function and even the nuclear location of U9 and U10 are unknown.[3]

Herpes virus codes for four specific snRNAs during infection of mammalian cells. The virus synthesizes four novel snRNAs, HSUR 1 to 4, of low abundance that are assembled with the snRNP core proteins and that acquire a typical trimethylguanosine cap.[48]

## C. Plant snRNAs

Fungal (*Neurospora* and *Aspergillus*) and pea plant snRNAs isolated with trimethylguanosine cap antibodies show remarkable homology to the snRNAs of higher animals.[49-53] U2 analogs were identified by hybrid selection with a cloned mammalian U2 gene and by immunoprecipitation with anti-*Sm* antibodies. After injection into *Xenopus* oocytes, the fungal and pea snRNAs were packaged into snRNPs and transported into the nucleus, indicating that the Xenopus snRNP proteins assemble with the injected snRNAs to form a particle capable of normal intracellular localization and transport. However, the efficiency of association with *Xenopus* snRNP proteins varied between species and the individual snRNAs.[49]

Homogenization of leaf tissue, followed by filtration and fractionation on Percoll cushions, prepares large quantities of purified plant nuclei and nucleoli. When low-molecular-weight RNAs are fractionated on gels, the overall pattern strongly

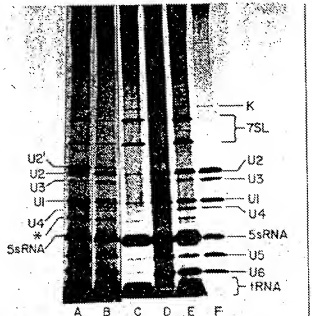


FIGURE 2. SnRNAs in the cytoplasm and nucleus of L929 cells. $2.5 \times 10^8$ $^{32}$P-Uridine-labeled L929 cells were pretreated for 10 min with 10 μg/ml cytochalasin B and enucleated on 12.5 to 25% Ficoll gradients. Cytoplasts and karyoplasts were harvested and the small RNA species were analyzed on 6 to 15% gradient gels. Lane A illustrates the small RNA species present in cytoplasts prepared from cells pulse labeled for 20 min with 80 μCi/ml $^{32}$P-uridine. Lane B is of the small RNAs present in an identical preparation of cytoplasts cultured for 60 min after enucleation and before harvesting. Lanes C and D display the stable RNAs in cytoplasts and karyoplasts, respectively, from cells labeled for 16 h with 5 μCi/ml $^{32}$P-uridine. Lanes E and F illustrate the stable RNAs in cells fractionated into cytoplasm and nucleus, respectively, using standard aqueous procedures. All lanes are from the same gel, although lanes A and B are from a longer exposure. (From Zieve, G. W., Sauterer, R. A., and Feeney, R. J., *J. Mol. Biol.*, 199, 259, 1988. With permission.)

resembles that of mammalian cells, and is similar among several higher plant species.[39] Sequencing of several of the plant RNAs comigrating with mammalian snRNAs confirm their identity as plant snRNAs. The sequence of broad bean U2 shows virtual identity at the 5′ end and strong homology at the 3′ end to mammalian U2, as well as a nearly identical secondary structure and a consensus Sm binding site in the same region as mammalian U2.[51]

Like mammalian U3, broad bean U3 is nucleolar in location, and shows about 50% sequence homology and essentially identical secondary structure homology with its mammalian counterpart.[50] Broad bean U6 is 80% homologous in sequence and virtually identical in secondary structure to rat U6, but it lacks seven nucleotides at the 5′ end, and has a somewhat different pattern of methylated nucleotides. Similar to the mammalian U6, bean U6 is hydrogen bonded to U4.[52,53] The cap of broad bean U6 appears to be a modified nucleotide of unknown nature, unlike that of mammalian U6, which is a methyl group on the gamma phosphate of the 5′ nucleotide.[54,55,56]

**Table 2**
***Saccharomyces cerevisiae* snRNAs[10,11]**

| SnR | Length | IP | Essential? | Comments |
|---|---|---|---|---|
| 2 | ~185 | No | ? | |
| 3 | 194 | No | No | Some U4 homology |
| 4 | ~192 | No | No | |
| 5 | ~193 | No | No | |
| 6 | ~110 | No | Yes† | Yeast U6 base paired to snR 14 |
| 7.1 | 214 | SM | Yes | Yeast U5 |
| 7.s | 179 | SM | Yes | 7.1 and 7.s are transcripts of same gene |
| 8 | 189 | No | No | |
| 9 | 188 | No | No | |
| 10 | 245 | No | No | |
| 11 | ~259 | No | ? | |
| 13 | ~125 | No | ? | |
| 14 | 159 | SM | Yes | Yeast U4 |
| 15 | ~158 | No | ? | |
| 16 | ~158 | No | ? | |
| 17 | 378 | U3 Weak SM | Yes | Yeast U3 |
| 18 | ~350 | No | ? | |
| 19 | 569 | SM | Yes | Yeast U1 |
| 20 | 1175 | SM | Yes | Yeast U2 Contains homology to U4, U5, U6 Interior 950 or not essential |
| 21 | ~165 | No | ? | Minor species |
| 22 | ~170 | No | ? | Minor species |
| 23 | ~190 | No | ? | Minor species |
| 24 | ~220 | No | ? | Minor species |
| 25 | ~265 | No | ? | Minor species |
| 26 | ~270 | No | ? | Minor species |
| 30 | 605 | Weak SM | Yes | Nucleolar |

*Note:* Essential snRNAs have been tested by gene deletion. SnRNAs marked ? have not been tested by gene deletion but are probably nonessential. SnR 4 is omitted because it is identical with snR3. There is no S.cR 12. All listed snRNAs are immunoprecipitable with trimethylguanosine cap antibodies, which is diagnostic for snRNAs, except for snR 6, which is precipitated only by its association with snR 14.

Data from References 10 and 11.

## D. Yeast snRNAs

In yeast cells, approximately 24 low-molecular-weight RNAs, ranging from 100 to 1175 nucleotides, have the 2,2,7 trimethylguanosine cap characteristic of the snRNAs[9,10] (Table 2). They are present in only 200 to 500 copies per nucleus, and are usually coded for by single copy genes instead of the multiple copy genes of higher eukaryotes. This correlates with the substantially fewer processed introns in yeast pre-mRNA.[34] Six of the yeast snRNAs are analogous in structure and functions to the major mammalian snRNAs U1 to U6 and are essential for cell viability[9,10] (see Table 3). These RNAs have the consensus sequence responsible for the

**Table 3**
**Essential snRNAs of *Saccharomyces cerevisiae***

| SnR | Length | IP | Spliceosome component | Functional analog | Ref. |
|---|---|---|---|---|---|
| 19 | 569 | SM | Yes | U1 | 58, 59 |
| 20 | 1175 | SM | Yes | U2 | 60 |
| 17A[a] | 328 | U3/SM | No | U3 | 57 |
| 17B | 328 | U3/SM | No | U3 | 57 |
| 14[b] | 159 | SM | Yes | U4 | 66 |
| 7[c] | 214 | SM | Yes | U5 | 68 |
| 7.5[c] | 179 | SM | Yes | U5 | 68 |
| 6[b] | ~110 | No | Yes | U6 | 65, 66 |

*Note:* Deletion of genes is lethal for all listed snRNAs, except SnR 17, which requires deletion of both SnR 17A and SnR 17B for lethal effect. Deletion of SnR 6 gene has not been tested.

[a] SnR 17a and 17b are highly homologous transcripts from two distinct genes. Deletion of both is required for lethal phenotype.

[b] SnR 14 and SnR 6 are base-paired to each other, similar to mammalian U4 and U6.

[c] SnR 7l and 7s are two different transcripts from the same gene.

binding of the Sm core proteins and are precipitable by anti-Sm sera.[11] Unlike mammalian cells, the nucleolar snRNA, yU3 (snR 17) has the Sm binding site.[56,57] The snRNAs from the yeast *Saccharomyces cerevisiae* are often larger than their mammalian counterparts, especially the yU1 (snR 19) which is 569 nucleotides compared with mammalian 165 nucleotides and the yU2 (snR 20), which is 1175 nucleotides compared with mammalian 180 nucleotide U2.[58-60] However, regions of these snRNAs are highly homologous to their mammalian counterparts.

Although yeast snRNAs are most thoroughly characterized in *S. cerevisiae*, this yeast may be unusual among fungi. The similar but less characterized yeast *Schizosaccharomyces pombe* has a U2 analog similar in size to the mammalian U2, and the pattern of snRNAs immunoprecipitated from *S. pombe* more closely resembles the immunoprecipitation pattern from other fungal or animal cells than that of *Saccharomyces cerevisiae*.[60] Although the studies reported here use *S. cerevisiae* as the model system, this species may be unique and not representative of most yeast or protist snRNAs.

Many of the essential snRNAs in *S. cerevisiae* have been identified as homologous to the major snRNAs in higher animal cells and are now referred to by the mammalian names. The yeast snRNAs U1, U2, U4/U6, and U5 function in pre-mRNA splicing with analogous functions to the mammalian snRNAs.[58,61,62] YU1 is larger than mammalian U1, and primary sequence homology is limited to scattered single-stranded regions near the 5′ end, including absolute conservation of the 11 nucleotides at the 5′ end involved in binding to the 5′ splice junction of introns.[58,59] YU2 is also larger than mammalian U2, but when most of the extra sequences are removed, the snRNA participates effectively in splicing.[63,64] This large RNA has a region highly homologous to metazoan U2, and large regions homologous to vertebrate U4, U5, and U6, but not to U1. There are also three separate Sm binding sites (one each in the U2, U4, and U5 homology regions). The "extra" nucleotides of both yU1 and yU2 snRNAs may allow extensive interactions between the two RNAs during the splicing process. There are 11 regions of complementarity between yU1 and yU2, ranging from 9 to 16 nucleotides, and found throughout the lengths of both molecules that could be involved in base pairing interactions with each other.[58,62,64] YU6 (snR 6) is over 80% homologous to mammalian U6, and, like the mammalian species, base- pairs with yU4 (snR 14).[65,66] One of the other yeast snRNAs (snR 7) shares homologies with the true U4, but is not essential for cell viability.[67] YU5 (SnR 7) has a 70-nucleotide sequence strongly homologous in secondary structure, though only weakly homologous in sequence, to mammalian U5.[68] Two forms of yU5, yU5 L, and yU5 S, are produced in equimolar amounts, with one about 35 nucleotides longer at the 3′ end than the other. Both have a trimethylguanosine cap and both are associated with yeast spliceosomes.[66]

Several yeast snRNAs have been identified hydrogen bonded

**Table 4**
**Protein Composition of the snRNPs**

| | | Molecular weight (kDa) | | | |
|---|---|---|---|---|---|
| snRNP | Protein | SDS-PAGE | cDNA | Ref. |
| Shared "core" | B′ | 29 | (Same as B?) | 157 |
| | B | 28 | 29.1 | 157 |
| | D | 16 | 13.3 | 160 |
| | D′ | 18 | — | 149a |
| | E | 13 | 11 | 162 |
| | F | 12 | | |
| | G | 11 | | |
| Neural-specific variant of B | N | 28 | 29.1 | 20 |
| U1 snRNP | "Core" | — | | |
| | 70 | 68—70 | 52 | 164a |
| | A | 33 | 31 | 132 |
| | C | 22 | 17.4 | 166 |
| U2 snRNP | "Core" | — | | |
| | A′ | 32 | | 169 |
| | B″ | 28 | 25.5 | 167 |
| U3 snRNP | | 74 | | |
| | | 59 | | |
| | | 36 | | 25 |
| | | 30 | | |
| | | 13 | | |
| | | 12.5 | | |
| U4/U6 snRNP | "Core" | — | | |
| | | (23) | | |
| | | (12) | | 153 |
| | | (10) | | |
| | PRP4 | 52 | 52 | 156a |
| U5 snRNP | "Core" | — | | |
| | | 25 | | 153 |
| | | 70—100 | | 152,154 |

Data are from Reference 148, except where noted.

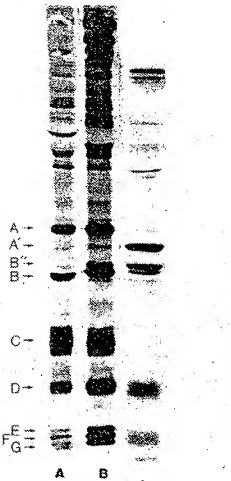


FIGURE 3. SnRNP proteins in the U1 and U2 snRNP particles L929 cells were labeled with 10 μCi/ml $^{35}$S-methionine (lanes A and B) or 10 μCi/ml 3H-leucine for 16 h (lane C). Nuclear fractions prepared by aqueous cell fractionation were immunoprecipitated with a (U1)RNP monoclonal antiserum (lane A), the Y12 Sm antiserum (lane B), or a U2 patient antiserum (lane C). The B, D, E, F, and G common core proteins and the U1-specific A and C' and the U2-specific A' and B'' proteins are indicated. The U2-specific A' protein is leucine rich and methionine poor and is most obvious in cells labeled with leucine. The 70-kDa U1-specific protein is present in substoichiometric amounts and is difficult to identify on one-dimensional gels. (From Feeney, R. E., Sauerer, R. A., Feeney, J. L., and Eleve, C. W., *J. Biol. Chem.*, 264, 5776, 1989. With permission.)

to preribosomal RNA in the nucleolus, including yU3.[56] The nucleolar snRNA yU3 is 50% longer than mammalian U3, but it shares nearly 50% homology in nucleotide sequence with the first 100 nucleotides of U3 and a virtually identical secondary structure, with the exception of an additional loop in the middle

**Table 5**
**Antibodies to the snRNPs**

| | Antigenic determinants | snRNPs precipitated | Ref. |
|---|---|---|---|
| Anti-Sm | Mainly B, B' and/or D, D' Rarely E or A | U1, U2, U4, U5, U6 U7, U8, U9, U10, U11 | 22, 121 |
| Anti-(U1)RNP | 70-kDa protein often A Occasionally C | U1 | 123, 124, 126, 127 |
| Anti-(U1, U2) | A (U1) and B'' (U2) proteins | U1, U2 | 24 |
| Anti-(U2) | A' and B'' proteins | U2 | 23 |
| Anti-(U3) | 34-kDa protein | U3 | 26 |
| Anti-La | 50-kDa protein | U1, U6 | 2, 185, 186 |
| Anti-m3G | 2, 2, 7 Trimethylguanosine | All except U6 | 140, 141 |
| Anti-m6A | $N^6$-Methyl-adenosine | U2, U4, U6 | 143 |

of the molecule where the "extra" nucleotides (relative to mammalian U3) are inserted.[57] SnRNPs containing yU3 and U3 also share at least one homologous protein, as yU3 is immunoprecipitated with a human anti-U3 antisera reactive to a 34-kDa protein.[57] Unlike mammalian U3 snRNPs, Sm antisera weakly precipitates yU3. A Sm consensus sequence is present in the extra sequences of the yeast snRNA.[58] Several of the nucleolar yeast snRNAs, snR4, snR5, snR8, snR9, do not have the snRNP core binding sequence, are not associated with Sm core proteins, and can be deleted from the cell with minimal effects on cell growth.[26] The 605-nucleotide yeast snRNA, snR30, is essential for cell growth, and its primary sequence has similarities to the other nucleolar snRNAs, suggesting that it may also function in the nucleolus.[59]

## E. SnRNA Genes

### 1. Overview and General Characteristics

The genes for all six major snRNAs have been cloned and studied in a variety of organisms. These genes are transcribed by RNA polymerase II, except for U6, which is transcribed by RNA polymerase III.[12-14] Although species-specific differences exist, the true snRNA genes are present in a few tens of copies each per genome in higher animals, with a tendency toward fewer copies of genes in lower organisms. This parallels a trend of decreasing copies of the snRNAs themselves in lower organisms. Depending on the organism and the snRNA, there are often hundreds of nontranscribed, frequently truncated pseudogenes in the genome.[10]

In the yeast *S. cerevisiae*, the snRNA genes are single-copy genes (yU3 is derived from two closely related genes), and there are as few as 200 to 500 copies of each snRNA per cell.[58] Unlike the situation in mammalian cells, the yeast snRNA genes have promoters that are similar to yeast mRNAs. There are TATA boxes upstream of the genes and accurate initiation is directed by nucleotides immediately upstream of the genes.[58] In mammalian

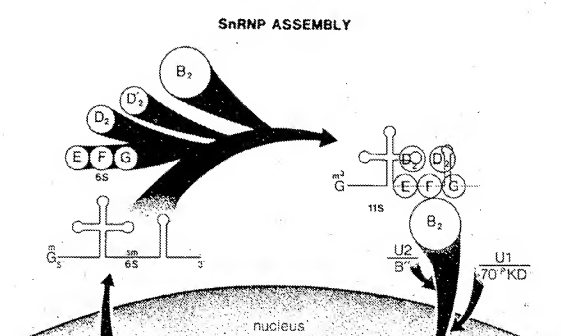


FIGURE 4. Cartoon model of the cytoplasmic assembly and nuclear exchange of the snRNP particles. Newly synthesized snRNPs appear transiently in the cytoplasm where they undergo 3' end processing and cap hypermethylation and assemble with the snRNP core proteins stored in large, partially assembled pools in the cytoplasm. The D, E, F, and G core proteins preassemble into a 6S RNA-free particle, and homooligomers of B, D, and D' protein are also present. The dynamic behavior of the U1 and U2 snRNP-specific proteins are also illustrated. There are large pools of the A and C U1 specific proteins and the A' U2-specific protein in the nucleus. However, there are no pools of the U2 specific B'' protein or the U1-specific 70-kDa phosphoprotein.

cells the control elements of the snRNA genes significantly differ from the control elements of mammalian mRNA genes. This may reflect the high rates of transcription required of these genes to produce the abundant snRNAs.

Studies of the structure of the different cloned mammalian snRNA genes reveal the snRNA genes lack a TATA box at -30 typical of mRNA genes, but have conserved 5' and 3' sequence elements that are necessary for proper initiation and termination (Figure 6). The 5' control elements include an enhancer at -200 to -250 nucleotides upstream, called the distal sequence element (DSE), and a proximal sequence element (PSE), typically located at -50 to -60 nucleotides upstream.[71-75] (See Reference 76 for a compilation of sequences and Figure 6 for a schematic representation.) The DSE contains a conserved octamer motif ATGCAAAT that is also homologous to a sequence found in SV 40 and immunoglobulin enhancers, and a SP1 binding site.[77] This element typically increases the rate of transcription 10- to 20-fold and is usually, but not always, orientation independent.[73] The PSE is not as conserved as the DSE and is functionally analogous to the TATA box directing transcription to begin at nucleotide 1.[75] The snRNA genes also have a short sequence called the 3' box located at about 10 to 30 nucleotides downstream of the coding sequence. This sequence, and also proper snRNA 3' sequences, are required for proper termination of transcription, suggesting interactions between the 5' and 3' end of the genes.[78-80] Surprisingly, there are homologies between the structure of the U6 genes and those of the other snRNAs, despite the fact they are transcribed by different RNA polymerases.[81-84] With the exception of U6, the snRNA genes are not accurately transcribed in the available mammalian *in vitro* systems that transcribe mRNA genes.[85] In addition to the obvious difference in sequence organization, this also suggests major differences in the transcription of the snRNA and mRNA genes. Recently, U1 genes were faithfully transcribed in isolated nuclei and in *in vitro* systems prepared from a highly concentrated extract of *Xenopus* oocytes or sea urchin blastula.[86-88]

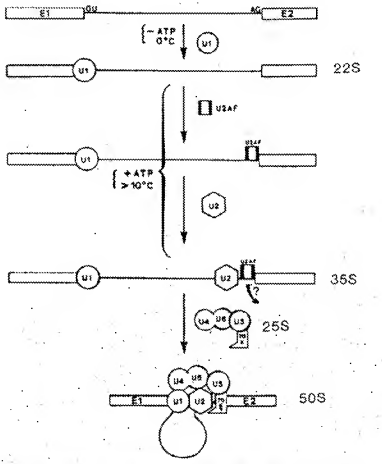


FIGURE 5. Channon model of the sequential steps in spliceosome assembly. U1 snRNP binds to the substrate in an ATP-independent fashion to generate a 22S intermediate. U2 snRNP adds after additional cofactors to form the 35S intermediate. A 25S particle of U4/U6 and U5 snRNPs then adds to form the mature functional 50 to 60S spliceosome. (From Ruskin, B., Zamore, P. P., and Green, M. R., *Cell*, 52, 207, 1988. With permission.)

The U1 and U2 genes often appear in clusters, and there are up to several hundred (exact numbers are not known) nontranscribed U1 pseudogenes.[79,89-91] Although there are pseudogenes for other snRNA genes, most of them are truncated and are usually not present in as great a number as for U1.[79] In *Xenopus* and sea urchin there are large tandem arrays of embryonic U1 genes that are transcribed at high rates in early development, in addition to several adult U1 genes that are present in less than ten copies each and are scattered throughout the genome.[92,93] The structures of the genes for the major snRNAs are reviewed below.

### 2. U1

Restriction mapping at high stringency suggests there are approximately 30 true human U1 genes in the human genome, but from 500 to 1000 pseudogenes.[79,94,95] Currently, cloned full-length U1 genes from human cells all have coding sequences identical to the canonical U1 snRNA sequence; however, two results suggest there may be additional bona fide U1 genes.[96,97] Sequence analysis of the U1 snRNAs in human cells reveals there is microheterogeneity in between 5 and 15% of the U1 snRNAs. These variants differ by from one to three nucleotide changes, and they occur in both single- and double-stranded regions of the snRNA.[96] The genes for these variants have not been identified and may be among the large family of U1 genes that were previously considered pseudogenes. Also, newly synthesized U1 snRNA is several nucleotides longer than mature U1, and the

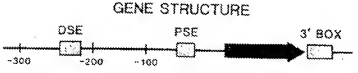


**FIGURE 6.** Canonn model of the regulatory elements of the snRNA genes. Regulatory elements of the snRNA genes differ from those of other genes transcribed by RNA polymerase II. The genes have a distal sequence element (DSE) that acts as an enhancer and a proximal sequence element (PSE) that acts as a promoter and is analogous to the TATA box of mRNA genes. A 3' box is essential for proper termination of transcription.

extra nucleotides are removed in the cytoplasm during snRNP assembly. Recent studies have shown that newly transcribed U1 precursors in human cells have heterogeneous 3' flanking sequence.[95] All previously sequenced human U1 genes have the identical 19 nucleotide sequence extending 3' from the coding sequence, suggesting that other variant, previously unsequenced genes are also transcribed. In support of this hypothesis, a cloned major U1 gene and a class 1 pseudogene hybridized to different subsets of the U1 transcripts, suggesting that at least some of the previously described U1 pseudogenes are actually transcribed.[97] This suggests the number of true U1 genes is greater than the 30 genes with the canonical U1 sequence and also include a small number of genes with variant coding sequences that also include variations in the 3' transcribed sequence.

Sequence analysis of the available cloned human U1 genes reveal homology between the genes in the 5' flanking regions that extend at least 2.6 kb upstream, while sequence conservation extends only about 50 base pairs at the 3' end.[72,94,95] None of these clones show a 5' sequence identical to the TATA box consensus common in other genes transcribed by RNA polymerase II. However, the initiation site is similar to other polymerase II genes, including starting with an adenosine. Alignment of the U1 sequences with other cloned snRNA genes identified two conserved sequence motifs, located at -51 and -212 base pairs, representing the PSE and the enhancer-like DSE, respectively.[72-75] Other, less conserved sequence motifs are also required for efficient transcription.[76,95]

Although highly conserved within a species, the 5' sequences are not homologous with those of other species except for the conserved proximal and enhancer elements, and show little or no homology with sequences flanking pseudogenes.[70] High stringency hybridization using probes to the 5' flanking regions of the U1 gene show that there are about 30 true genes for human U1, and a comparable number for rat and mouse.[94]

A conserved region at the downstream end of the U1 gene is necessary for proper 3' end formation, and its deletion results in the formation of much longer U1 transcripts.[78-80] The region required, called the 3' box, is from 9 to 28 nucleotides 3' of the coding sequence for human U1 genes. Surprisingly, the actual U1 coding sequence is apparently not required for accurate 3' end formation, as its deletion and replacement with foreign sequences, such as parts of globin genes, does not affect 3' end formation.[78] However, a functional U1 or U2 PSE is an absolute requirement for formation of the proper 3' end.[78,80] Replacement of the snRNA promoters with mRNA promoters results in polyadenylation of an extended 3' end, apparently due to a polyadenylation sequence located downstream of the normal 3' end.[80] Hernandez and Weiner[80] also found that the U1 or U2 5' enhancer was required for accurate 3' end formation, though this was disputed by other labs,[80] who used more or less intact U1 sequences rather than foreign gene sequences in their genetic constructs. Taken together, the data suggest that the U-RNA transcription complex involves factors that recognize both 5' and 3' sequences that are not found in other RNA polymerase II transcription complexes, and that these factors are required to form the unique 3' ends of snRNAs.

Single-point mutations within the 3' box[100] of U1 inhibits proper 3' end formation in both *Xenopus* oocytes and HeLa cells, as does single-point deletions or insertions. The degree of correct 3' end formation in these mutant genes varies from 80% to only 22% of wild-type controls. In general, HeLa cells process the human U1 3' ends more accurately than oocytes when injected or transfected with the same mutant U1, and double-point mutations have a greater inhibitory effect, though all mutations tested showed at least partial ability to generate proper 3' ends. The 3' box is apparently functionally equivalent between different snRNAs and species, since substitution of the normal U1 3' box with a rat U3 3' box allowed nearly normal 3' processing, with the only difference being a more extended 3' transcript.[100]

Using low-stringency hybridizations, there are between 500 and 1000 human U1 pseudogenes.[70] These pseudogenes fall into three distinct classes: class I, which has scattered base substitutions in the U1 sequence itself and considerable flanking homology with the true genes; class II, which has base substitutions and is truncated at the 3' end, with little homology to the true genes in the flanking regions; and class III, which is full length in the U1 region itself, and has flanking regions different from either class II pseudogenes or true genes. It is proposed that class I pseudogenes are generated by gene duplication, because of the conservation of the flanking regions, while class II and III

pseudogenes are created by a RNA-mediated insertional mechanism. The microheterogeneity observed in the U1 snRNA discussed above suggests some of the class I pseudogenes are actually transcribed and assembled into mature nuclear U1 snRNPs.[96,97]

In humans, the U1 genes are clustered on a single band (though it is possible that a few genes could be located elsewhere), 1p36.3, located toward the tip of chromosome 1, with at least 20 kb separating individual genes.[98] The class I pseudogenes are clustered in another band on chromosome 1.[99]

In *Xenopus* there are at least seven different U1 snRNAS, including two found in embryonic cells and five found in adult tissue.[92,101] The two embryonic U1 genes, xU1ba and xU1b2, are present in 500 copies each and are paired in 1.8 kb repeat units organized in large tandem arrays. The adult genes are present in approximately ten copies each scattered throughout the genome.[92] There appear to be relatively few pseudogenes in *Xenopus*, since microinjection of 20 randomly subcloned *Xenopus* U1 genes into oocytes were all transcribed, and this would be unlikely if there were more U1 pseudogenes than true genes in *Xenopus*.[101] The structure of *Xenopus* and human genes is similar, including a PSE at about -50 nucleotides and an enhancer-like DSE at -203 nucleotides that is necessary for normal transcriptional levels.[101]

SnRNAs are not accurately transcribed in *in vitro* systems prepared from mammalian cells. In a HeLa cell extract, a cloned U1 gene is transcribed by RNA polymerase II, but transcription begins 183 base pairs upstream of the 5' end of the actual U1 gene, and a sequence or sequences between -106 and 393 base pairs upstream (5') is necessary for transcription.[85] However, cloned snRNA genes are accurately transcribed when injected into *Xenopus* oocytes or when incubated with a concentrated extract of the *Xenopus* oocytes or sea urchin blastula.[86,87]

Unlike the vertebrate snRNA genes, the yeast U1 genes sequence elements are analogous to those of mRNA genes. Both yU1 and yU2 have a TATA box located at nucleotide -98, and the sequences immediately upstream of transcription initiation are virtually identical between the two snRNAs.[38,39]

### 3. U2 Genes

The structure of the human U2 genes has many similarities to the structure of the U1 genes, including 5' distal and proximal elements and 3' sequence elements that are necessary for proper transcription.[74,75] However, there are far fewer U2 pseudogenes than U1 pseudogenes, and the U2 pseudogenes in humans are usually severely truncated at their 3' ends.[90] Also, the U2 genes are arranged in tandem arrays with a repeat length of 6 kb located on a band in chromosome 17.[90,101,102]

Mutation of the 5' and 3' sequence elements of the U2 genes and analyses of their effects on transcription have helped identify common regulatory features among the snRNA genes.[72-75] Like U1, U2 has no true TATA box, but requires both a DSE at -220 and a PSE at -50 for transcription. Deletion of the DSE (enhancer) between -218 and -295 base pairs causes a decline in the level of transcription to less than 5% of control levels, although accurate initiation still occurs. Deletion of the 5' sequences to the -3 position, removing both the DSE and proximal element, completely abolishes U2 transcription, while deletion of the first 92 base pairs 3' to the U2 coding sequence has no effect on transcription.[74] Inversion of the DSE reduces accurate transcription of U2 to less than 5% of control levels, but transcripts initiated at abnormal upstream sites are stimulated.[75] Accurate transcription occurs at control levels independent of location upstream (at least from -61 to -198 base pairs), provided the DSE sequence is in its proper orientation. Deletion of four base pairs from the DSE between -218 and -223 abolishes transcriptional activation to the same extent as deletion of the entire sequence. The DSE has many of the characteristics of typical enhancers, such as position-independent activation of transcription, but also provides information necessary for proper choice of initiation site.[73,75,103]

There are no detectable developmental variants of the U2 genes in *Xenopus*. The *Xenopus* U2 genes are arranged in tandem arrays with an 831-bp repeat length, containing all the signals necessary for transcription.[74,102] There are about 500 copies of the U2 genes in *Xenopus*, with no identified pseudogenes. The control of U2 gene transcription in *Xenopus* is similar to that described in humans, except the enhancer-like DSE, located at about -260 base pairs, is completely orientation-independent.[74] Sequence analysis of this element shows strong conservation with the human U2 DSE, and is homologous to the SV-40 and immunoglobulin heavy gene enhancers. Deletion of the DSE reduces transcription to only about 5% of control levels, while deletion of the proximal element abolishes transcription entirely.[74]

Recent studies have isolated and characterized the yeast U2 gene (snR 20). An oligonucleotide probe for human U2 RNA isolated a normal-sized U2 gene from *Schizosaccharomyces pombe*, but bound to an 1175-nucleotide-long RNA in *Saccharomyces cerevisiae*. The gene itself, LSR1, has sequence elements similar to other yeast snRNA genes, including a TATA-like box at -98 nucleotides and a 3' thymidine-rich region. Deletion of this gene in yeast results in loss of viability and cell death.[40]

### 4. U3, U4, U5, and U7

The genes for U3 and U4 are cloned, but are not as well characterized as the genes for U1 and U2.[104,105] Like U1 and U2, the U3 and U4 snRNA genes are dispersed in the genome and there are abundant pseudogenes for each. Crude estimates of the copy number for these other genes suggests that there are a few hundred genes and pseudogenes for each gene family, but only a few tens of true genes for each snRNA in mammalian cells.[104,105]

The human U3 gene has been cloned and is present only once on a 2.3-kb fragment, indicating that if the human U3 genes are in tandem arrays, they are separated by at least a kilobase.[106] The U3 gene has regions homologous to the DSE, PSE, and 3' processing signal of U1 and U2, as well as a sequence adjacent to the DSE that is specific to U3 and conserved in U3 genes from

other species. Southern hybridizations show that between six and nine copies of the human U3 gene exist, and that the genes are located in the nuclear, rather than the nucleolar, DNA.[104] The U3 genes are localized in the nucleoplasm, not in the nucleolus like the mature U3 snRNP.[106]

Mammalian U4 genes are in three families with minor differences in coding sequences.[105] Clones of two of the genes, U4b and U4c, are on a single 930-nucleotide fragment; however, slight differences in the coding sequence of the U4b gene and the known U4b sequence suggest that the cloned gene may be a minor variant and not the main gene cluster.[105] Chicken cells also have two closely linked U4 variants.[107] At least 80 U4 genes exist in human cells, and there are many truncated U4 pseudogenes as well.[105] The U4 genes have 5' sequence elements common to other snRNA genes, including the PSE at -50 and an enhancer-like DSE at -210. In addition, there is an element centered around -140 that is homologous between the U4 gene families, but absent from U2 genes, and may represent a control element unique to U4.[105] In *Xenopus*, at least three different U4 snRNAs were identified, and the xU4b is preferentially transcribed during oogenesis in contrast to the other two that are transcribed in adult tissue.[108]

U5 genes in *Xenopus* are present in several arrangements, including approximately 100 genes found in 583 bp repeats.[109] The U5 genes have the required distal and proximal sequence elements similar to those described for U1 and U2. The octamer motif found in the DSE is in the opposite orientation to that found in U1 or U2, and an additional required sequence element was found in the DSE that is not present in U1 or U2. The U5 gene is expressed more efficiently than the U2 gene when injected into *Xenopus* oocytes.[109]

Unlike other yeast snRNA genes, the yeast U3 gene (snR 17) has two functional genes, each with 328-nucleotide-long coding sequences that share 96% homology.[56,57] The flanking sequences are much less conserved. Both genes are transcribed, with yU3a at a tenfold higher level than yU3b. The two RNAs are identical in localization, immunoprecipitability, and base pairing to pre-rRNA. Deletion of both genes is lethal; however, either one of the genes can be deleted from the cell without effect. Transcription of these two genes is coordinately regulated, since production of snR 17b increases when snR 17a is deleted.

The yeast yU3 genes have a TATA box 85 nucleotides upstream from the initiation site.[56,57] The exact nucleotide sequence upstream of the initiation site is preceded by $A_{10}GT$ and differs from that of the spliceosomal yeast snRNAs which have stretches of Cs, As, and Ts.[56]

Five U7 genes were identified in a 9.3-kb cluster in the sea urchin *Psammechinus miliaris*. Only one of the genes, U71, has the identical sequence as the cloned cellular U7 snRNA, but it is possible the others are transcribed at low levels. The U71 gene has 5' DSE and PSE and a 3' terminal box similar to that of mammalian snRNAs.[110]

### 5. U6 Genes

Despite abundant evidence that the U6 snRNA is a RNA polymerase III transcript, the genes for U6 share many features common to the other snRNA genes, which are polymerase II transcripts.[13,14] In fact, the U6 genes injected into *Xenopus* oocytes are transcribed by both RNA polymerase II and III; however, the polymerase II transcripts are aberrant and their significance is unknown.[111] Similarities between the U6 gene and other snRNA genes include a required sequence at -43 to -67 nucleotides that is homologous to the U1 and U2 PSE, as well as an enhancer-like DSE at least 245 nucleotides upstream.[81,84] Deletion of the DSE decreases transcription of the microinjected gene in *Xenopus* by over 90%.[81] The U6 DSE is homologous to other snRNP DSEs, and it can replace the DSE on a U2 gene, a known polymerase II transcript. The DSE binds a factor common to the U2 DSE (and presumably other snRNA genes, all of which have the DSE and octamer motif).[77] Excess U2 genes coinjected into *Xenopus* oocytes with U6 genes reduces transcription of the U6 genes, as does a synthesized oligonucleotide containing the octamer motif, but competition is abolished if the competing U2 genes have the DSE deleted.[77] Sequences upstream of the DSE are also involved in transcription, since deletion of regions upstream of -280 nucleotides reduce transcription of *Xenopus* U6 genes to half the control levels.[77] A sequence similar to the TATA box found upstream of mRNA genes is present upstream of the U6 gene, but deletion of this sequence has only minor effects on U6 transcription.[77]

U6 genes also share an internal region at 48 nucleotides that is homologous to internal control regions of other polymerase III transcripts, such as 5S rRNA.[112] However, this internal control region, the U6 A box, can be deleted from U6 without affecting transcription. This region, however, competes with 5S genes for a common factor. When coinjected into oocytes, excess 5S genes compete with wild-type U6 to reduce or abolish transcription, and the U6 A box deletion mutant is much less affected than wild-type U6 in competition experiments.[77]

The U6 genes are randomly repeated in two large clusters in amphibians. One cluster has a 1.6-kb repeat and is present in 200 copies, and another has a 1-kb repeat present in about 500 copies per genome. The significance of the two different clusters is not known.[81]

In *Drosophila*, there are three U6 genes clustered on a single, 2-kb fragment.[112] Although the coding sequences for all three U6 genes are identical, the flanking regions, especially the 3' sequences, have diverged considerably. The *Drosophila* U6 genes have an upstream TATA box, but not the PSE and DSE found upstream of the mammalian U6 and other snRNAs. The *Drosophila* U6 genes also have an internal element (the A box) from 48 to 72 nucleotides that is homologous to control regions for tRNA and other polymerase III transcripts. Unlike most snRNA genes, transcription of the *Drosophila* U6 genes is species specific, since the *Drosophila* genes are accurately transcribed in *in vitro*

extracts using *Drosophila* nuclear extracts, but not if mammalian nuclear extracts are used. Likewise, mouse U6 genes cannot be transcribed using *Drosophila* nuclear extracts *in vitro*, but are transcribed using extracts from mammalian nuclei.[112]

In the fission yeast *Schizosaccharomyces pombe*, the U6 gene includes a transcribed intron.[113] The intron is in the region of the snRNA thought to be involved in the U4:U6 interaction. The intron is removed during processing in the nucleus and the mature U6 has a 77% homology with mammalian U6 snRNA. This is the first report of a snRNP involved in processing its own transcript.

## F. Isoforms and Developmental Regulation

A growing number of snRNA species include low-abundance variants with minor sequence heterogeneity. The functional significance of these variants and their possible contribution to differential RNA processing, including alternative splice site or polyadenylation site selection, is not known.[114] Developmentally regulated isoforms of U1 exist in some species. However, in most species, including human, chicken, and *Drosophila*, there are only one or two major forms of U1.[115] However, as discussed above, in human cells there is a small fraction of U1 that shows minor sequence differences with the major U1 snRNA. However, these variants showed no tissue- or cell type-specific distributions.[86]

Several other species express at least two distinct types of U1, generally differing by substitution of a few bases and differences in methylation of bases. Mouse cells contain two major isoforms of U1, U1a and U1b, which differ in six nucleotides, as well as one or two minor variants of U1a and U1b, all of which are developmentally regulated.[116,117] In the oocyte, greater than 85% of the U1 is U1a. Embryonic or fetal cells express U1a and U1b in equal quantities, and expression returns to almost totally U1a by a week after birth. Some U1b is present, even in adults, in tissues that have many undifferentiated stem cells, such as testes, spleen, or thymus.[116] Mouse Friend cells infected with spleen focus-forming virus (SFFV) express a minor U1a variant and four U1b variants not found in uninfected cells.[115]

In *Xenopus*, there are two embryonic U1 genes, xU1b1 and xU1b2, organized in large tandem repeats that are transcribed at a high rate in early previtellogenic oocytes and in early embryos.[92,108,118-120] Both adult U1 snRNAs U1a(1-5) and the embryonic snRNAs are switched on at the midblastula transition, but embryonic synthesis is soon switched off and only adult forms accumulate in the adult tissues.[108,118] The embryonic U1 genes accumulate in a 1:1 ratio in oocytes; however, during the short period of synthesis in early development, they accumulate in a ratio of 5 to 1, U1b1 to U1b2. This ratio of transcription is reflected when the genes are injected into unfertilized *Xenopus* eggs.[120]

The significance of the developmental regulation of U1 genes is unknown. The differences in sequence among the developmentally regulated U1 variants are not found in the 5′ end that forms the 5′ splice junction recognition sequence, but some of the base substitutions may alter the stability of the stem structure of the molecule with possible effects on snRNP assembly.[91] There is also differential accumulation of U4 isoforms during early *Xenopus* development.[108]

## G. Antibodies to snRNPs

### 1. Antibodies to snRNP Proteins

The discovery that antisera from patients with autoimmune diseases, such as SLE and mixed connective tissue disease (MCTD), recognize the snRNP particles provided a powerful new approach for studying the snRNPs.[75] At least five distinct specificities that recognize protein components of the snRNPs particles have been characterized from patients with autoimmune disease (Table 5).

Two major specificities of SLE autoimmune sera, Sm and (U1)RNP, present in over 40% of the patients with SLE, recognize protein components of the snRNP particles[121] (Figure 3). Although these were recognized clinically as antinuclear antibodies (ANA) for over 10 years, Lerner and Steitz[32] were the first to identify their reactive autoantigens as protein components of the snRNP particles. The autoantibodies directed against the Sm antigen are diagnostic of SLE, and anti-Sm antibodies immunoprecipitate the five major snRNAs U1, U2, U4, U5, and U6, and many minor snRNAs. SnRNA U3 is not precipitated. The anti-Sm sera recognize determinants on the common core of snRNP proteins, and the precipitation of the five different RNA components by a single antibody specificity was the first indication that the different snRNP particles share identical or at least similar proteins. (U1)RNP antiserum recognizes only the U1 snRNP, and it recognizes determinants on the U1-specific proteins.

The MRL strain of mice develops a SLE-like autoimmune disease with autoantibodies similar to those found in human SLE. When spleen cells from these mice are fused with myeloma cells they produce hybridomas secreting specific SLE autoantibodies. A variety of SLE hybridomas are now available, including many that produce monoclonal antibodies of Sm or (U1)RNP serotypes.[91,20,122-125] These antibodies have the advantage of unique specificity, avoiding the problems of mixed specificities and polyclonal responses that commonly appear in patient sera.

Detailed immunoblot characterization of various patient sera and monoclonal antisera show that the most common anti-Sm sera recognize the B protein, with occasional reaction against the D and E proteins, and rare activity against the U1-specific A protein (Figure 7). Anti-(U1)RNP sera react mainly with the U1-specific, 70 kDa protein, less often with the 33-kDa A, and rarely with the 22-kDa C, U1-specific proteins in Western blots.[126-128] Polyclonal antisera generated against a 23-amino acid, proline-rich sequence near the C terminus of the neural-specific B protein variant, N, recognize the B, N, and A proteins in immunoblots and define a potential Sm epitope.[129] This serum does not recognize the D protein, and this supports the conclusions from studies with the monoclonal antisera that there are multiple Sm epitopes.[122-125a] Anti-(U1)RNP antisera will precipitate the 70-

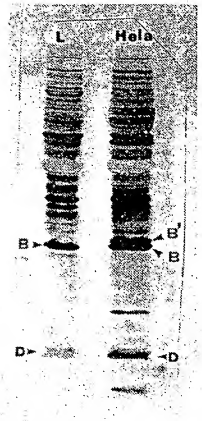


FIGURE 7. Western blot of human and rodent cell proteins with Sm antiserum. Whole cell fractions of murine L929 cells (L) and human HeLa cell line were analyzed on 13% polyacrylamide gels, transferred to nitrocellulose and probed with a human SM antiserum using an alkaline phosphatase coupled second antibody. The sera recognize the B and D protein in the murine cells and the B, B' and D proteins in human cells. (From Sauterer, R. A., Feeney, R. J., and Zieve, G. W., *Exp. Cell Res.*, 176, 364, 1988. With permission.)

kDa protein and the A and C proteins, either in association with the U1 snRNA or as free proteins, but will not recognize deproteinized U1 RNA.[126] Anti-Sm patient sera nearly always have considerable anti-(U1)RNP specificity as well, and anti-(U1)RNP patient sera often contain low titers of anti-Sm.[128] The high specificity of the monoclonal Sm and (U1)RNP sera is thus especially valuable when pure Sm or (U1)RNP antisera are needed for immunological analysis.

An anti-U2 specificity was discovered in the serum of a patient with scleroderma-polymyositis overlap syndrome.[22,126,131] These sera react with the 29-kDa B'' and 32-kDa A', U2-specific proteins immunoblots and specifically precipitate the U2 snRNP (Figure 3). Both monoclonal and polyclonal sera have been identified that cross-react with the U1-specific A proteins and the U2-specific B'' protein.[29,131] Sequence data available from cDNA clones indicate that the A and B'' proteins share large regions of identical sequence, which may explain the immunological cross-reactivity between the two proteins.[132] A U3-specific sera, reacting with the 34-kDa protein nucleolar protein fillagrin, was isolated from a scleroderma patient and has provided the first characterization of the U3 snRNP.[22,28,133,134] This antibody is now available as a monoclonal antibody.[28]

The B/B', D, D' 70 kDa, A, and B'', proteins, recognized by one or more of the existing patient sera, are probably the most immunogenic of the snRNP proteins. Injection of U1 purified by trimethylguanosine affinity columns and ion-exchange chromatography into mice resulted in monoclonal antibodies with Sm, (U1)RNP, or U1/U2 specificity, and none with specificities against other snRNP proteins, such as the core proteins E, F, or G.[135] These antibodies elicited in mice also competed effectively with patient antibodies of the same specificity, indicating that the antigenic determinants in both patient antisera and the mouse monoclonal antibodies were identical. This is consistent with the notion that the antibodies in autoimmune diseases are produced by an immune response directed against the snRNPs themselves, and not against viral or other foreign antigens that cross-react with snRNP proteins.[135,136] However, anti-(U1)RNP sera that recognize the 70-kDa, U1-specific protein cross-react with the retroviral-specific *gag* protein, which suggests the autoimmune response could be initiated by cross-reactivity with a protein from a viral infection.[137,138]

Small quantities of monospecific antibodies directed against specific snRNP proteins can be prepared using an affinity purification procedure by elution of antibodies bound to specific snRNP proteins on nitrocellulose blots.[131,139] This approach, although limited in yield, can provide specific probes to individual snRNP proteins of interest.

### 2. Antibodies Against snRNA Nucleotides

The 5' trimethylguanosine cap shared by all the snRNAs except U6 is another unique determinant common to the snRNP particles. Several labs have successfully raised both polyclonal and monoclonal antibodies to the trimethylguanosine cap.[140-145] These antibodies can be used to isolate either deproteinized snRNAs or intact snRNP particles by immunoaffinity techniques.[140-143] These antisera are now a preferred method for isolating both the major and minor snRNP particles. Although the U6 snRNA lacks the trimethylguanosine cap, it is immunoprecipitated by the cap antisera because it is associated with U4 in a single snRNP.[7,119] Unlike previous immunoaffinity approaches, such as anti-Sm or anti-(U1)RNP affinity columns, which require elution with high salt or denaturing agents, snRNPs bound to an anti-trimethylguanosine ($_{3}$G) affinity column can be eluted by addition of excess trimethylguanosine, and are thus isolated in a native, presumably functional form.[142,143]

Antibodies specific for $N^6$ methyladenosine (m6A), a modified nucleotide found only on U2, U4, and U6, were raised in rabbits, and found to immunoprecipitate U2, U4, and U6 quan-

titatively from cell extracts.[144] Some U1 is also coprecipitated using the m6A antibodies when cell extracts are used, but purified U1 does not react to the antibodies, and the precipitation of U1 occurs only in extracts containing U2. This indicates that a fraction (about 10%) of the U1 snRNPs interact with U2, confirming previous studies using alternative techniques.[145]

## H. Protein Composition of snRNPs

### 1. Protein Composition

Both biochemical isolation procedures and immunoaffinity techniques have allowed a detailed analysis of the protein composition of the snRNP particles (Table 4). With limited exceptions, the individual snRNP particles share a common core of six low-molecular-weight snRNP proteins, in addition to one to three snRNP-specific proteins (Figure 3). Different subsets of the snRNP proteins remain assembled with the snRNA, depending on the stringency of the isolation conditions. The snRNP-specific proteins are most easily removed from the particles and the D, E, F, and G core proteins are the most firmly associated with the snRNA. The major snRNP proteins are quite basic and can be resolved on two-dimensional gels using nonequilibrium pH gradient electrophoresis in the first dimension[17] (Figure 8).

By a combination of gel filtration and ion-exchange chromatography steps, several laboratories have produced fractions of highly purified U1, U1 and U2, and U2 to U6 snRNPs.[146-148] Except for slight differences in molecular weights (this report will use the molecular weights of Hinterberger et al.[146]), the results are in agreement.

The core proteins shared by U1, U2, U4, U5, and U6 in human cells are the 29-kDa B′, 28-kDa B, 18-kDa D′, 16-kDa D, 13-kDa E, 12-kDa F, and 11-kDa G proteins (Figures 3 and 8). Stoichiometric analysis based on the correction of amino acid labeling for known amino acid composition suggests these proteins exist in a particle with a stoichiometry of $B_1B'_1D_2EFG$.[17,148a] The proteins of U3 are different than those of the other snRNPs, and it does not share the proteins common to other snRNPs.[21,26] The B and B′ proteins are a closely spaced doublet in human cells, and the B′ protein is lacking in most nonprimate mammals (Figure 7).[17] Peptide analysis suggests that B′ is closely related to the B protein.[149] In SDS polyacrylamide gel electrophoresis the D′ protein often comigrates with the D protein; however, if urea is added to the separating gel, the D′ protein migrates with a higher molecular weight and is distinct from the D protein.[148,148a] At high ionic strength only the D, E, F, and G proteins remain associated with the snRNA, and they protect the Sm consensus sequence from RNAase digestion in a pattern similar to that with particles that also contain the B protein.[37] Also, particles are never found that contain the B protein, but not the D, E, and F proteins. This suggests that the B protein assembles with the snRNP through contacts with the other snRNP particles and not by direct binding to snRNA. This is consistent with the assembly studies discussed below, which demonstrate that the D, E, F, and G particles initially assemble into a 6S core that binds to the snRNA followed by two copies of the B protein.[17,150,151] In human cells the two copies of the B protein are replaced by one copy each of B and B′. Binding of the snRNP core proteins is directed by a sequence motif of PuA(U)nGPu, which is found in a single-stranded region in the 3′ half of all the anti-Sm precipitable snRNP particles (with the exception of U6)[28,37] (Figure 1). This sequence is necessary and sufficient for core protein binding. Insertion of this sequence into a single-stranded region of a heterologous RNA will direct assembly of the snRNP core proteins.[28]

In addition to the core proteins, each snRNP has one or more unique proteins. The unique proteins of U1 are clearly established because of the ability to isolate pure fractions of U1 snRNP by both biochemical and immunoaffinity procedures.[144,146,147] The U1-specific proteins are the 70-kDa protein, the 33-kDa A, and the 22-kDa C proteins (Figures 3 and 9). Stoichiometric studies suggest there are two copies each of the A and C proteins in each U1 snRNP, in addition to the common core, and that the 70-kDa protein associates with some, but not all, U1 snRNPs[17] (Figure 9). The 70-kDa protein is the only phosphoprotein in the U1 snRNP.[40]

Purification of the U2 snRNPs using either biochemical or immunological approaches identify two unique proteins, a 32-kDa methionine-poor protein designated A′ and a 28-kDa protein ($B''$) that are both immunoreactive with anti-U2 autoimmune sera[53,152,153] (Figure 3). As discussed earlier, the U2-specific $B''$ protein is homologous to the U1-specific A protein, suggesting that they may be the products of a gene-duplication event.[131]

Isolated U5 snRNPs include the common core polypeptides, a 100-kDa protein that degrades to a 70-kDa polypeptide, and a 25-kDa protein.[152-154] Several results suggest the 100-kDa protein may bind to the 3′ splice site.[152,154] The 100-kDa protein recognizes immobilized 3′ splice site sequences, and binding can be competitively inhibited by a 19-nucleotide fragment spanning the splice junction. Also, the protein binds normal, but not mutant, 3′ splice junctions, and can be immunoprecipitated from snRNP extracts with anti-Sm or anti-trimethylguanosine antibodies in complexes that include RNAase-protected, 3′ splice junction sequences. The protein cannot be immunoprecipitated by U1-specific antibodies, but coprurifies with biochemical fractions containing U5 snRNPs, in low (1 mM) $Mg^{++}$.[152,154] High $Mg^{++}$, which stabilize most snRNPs, disrupts the interaction of the 100-kDa protein with the core U5 particle.[152,155] In yeast, a 260-kDa protein was identified that binds to the U5 snRNP. The protein is also found in the ATP-dependent U5/U4/U6 complex that appears during pre-mRNA splicing.[156] The prp4 gene in yeast, originally identified as an essential gene for pre-mRNA splicing, has recently been identified as coding for a 52-kDa protein that is a component of the U4/U6 snRNP.[156a]

Although the shared core of snRNP proteins remains associated with the snRNAs under rigorous isolation conditions, the association of the other species-specific snRNP proteins is sensitive to ionic strength. In HeLa cell extracts made in 0.5 *M* NaCl, there are no specific proteins other than core proteins isolated

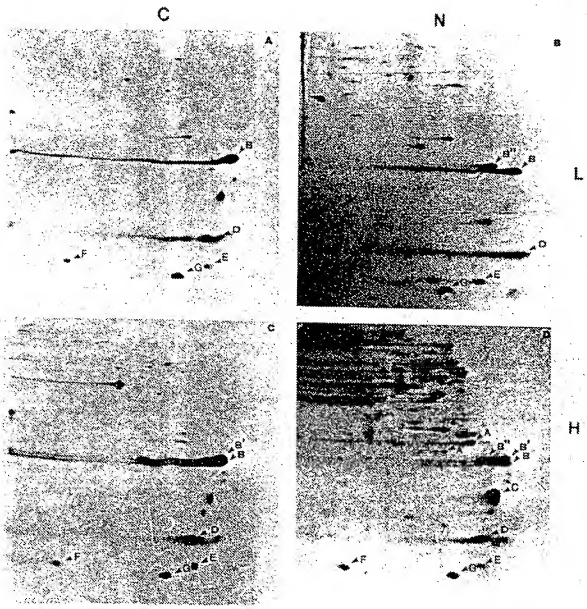


**FIGURE 8.** Two-dimensional gel electrophoresis of Sm immunoprecipitates of cytoplasmic and nuclear fractions from HeLa and L929 cells. HeLa (H) and L929 (L) were labeled 4 h with $^{35}$S-methionine, and cytoplasmic (C) and nuclear (N) fractions were immunoprecipitated with Sm antisera by standard procedures and analyzed by two-dimensional gel electrophoresis using non-equilibrium pH gradient gel electrophoresis in the first dimension. Cytoplasmic and nuclear fractions from L929 cells (A and B, respectively) and from HeLa cells (C and D, respectively) are illustrated, and major snRNP proteins are marked. Note the presence of B' in the cytoplasm and nucleus of HeLa cells but not in L929 cells. (From Feeney, R. J., Sauerer, R. A., Feeney, J. L., and Zieve, G. W., *J. Biol. Chem.*, 264, 3776, 1989. With permission.)

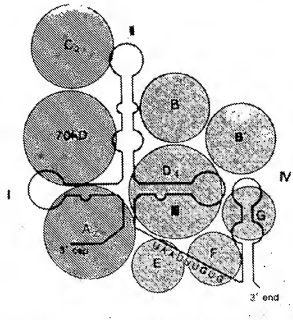


**FIGURE 9.** Suggested structure and stoichiometry of the U1 snRNP particle. The proteins are positioned based on structural data discussed in the text. The cross-sections of each protein subunit are from spherical particles where the volume is proportional to the molecular weight. New data suggest the D4 particle is actually $D_3$ $D'_1$.

with the U4/U6 and U5 snRNPs.[145] Also, high magnesium, which stabilizes the U1- and U2-snRNPs and keeps the U1 and U2 specific proteins associated with the snRNPs even in 1 *M* CsCl, dissociates the U5 100-kDa protein.[153,154,155]

A U3-specific antiserum isolated from a scleroderma patient and reactive with the 34-kDa protein fibrillarin on Western blots, allowed the first identification of the U3 snRNP proteins.[21,76] Two phosphorylated proteins of 74 and 59 kDa, and four non-phosphorylated proteins of 34, 30, 13, and 12.5 kDa are specifically immunoprecipitated in HeLa cell extracts.[21]

### 2. cDNA Clones of Individual snRNP Proteins

cDNA clones for the B, N, D, and E snRNP core proteins and the A, C and 70-kDa, U1-specific proteins and the A' and B" U2-specific proteins are now identified and sequenced.[20,156,157,158,159,160] The protein sequence information has helped identify functional domains in the proteins. The 70-kDa and A, U1-specific and the B", U2-specific proteins, but none of the sequenced-core snRNP proteins, have "RNP domains" found in a large number of RNA binding proteins and are suggested to make direct contact with the RNA.[156] The availability of clones has also helped define the antigenic domains of the proteins, and sequence data has allowed stoichiometric analyses based on amino acid labeling corrected for known amino acid composition.[17,129] Clones have been isolated by using either expression libraries where cDNAs were identified by their ability to direct the synthesis of immunogenic polypeptides, or by oligonucleotide probes generated from partial protein sequences.

#### A, B, N, D, AND E SNRNP CORE PROTEINS

cDNA clones coding for the B protein predict a 29.1-kDa protein that is unusually rich in glycine and proline residues.[157] Although the clones produce only a single polypeptide when analyzed by *in vitro* translation, hybrid selection using the clone as a probe identifies a mRNA or mRNAs that translate both the B and B' proteins. This suggests that the B and B' proteins are derived from either closely related genes or alternative splicing of a single pre-mRNA. A neural-specific form of the B protein, the N protein, has also be identified.[20,129,158] The B and N proteins have identical protein sequences.[20] The sequences for B and N predict proteins with a series of proline-rich repeating units in the C terminus, and a short, proline-rich region homologous to regions in the U1-specific A, C, and the hnRNP C proteins, which may be involved in RNA binding.[20] Homology to the D protein is limited to short aligned regions.[158] The N protein is of identical protein sequence in rats and humans.[20] This is a remarkable degree of conservation and suggests that all regions of the protein are essential for its function.

A genomic clone for the D protein derived from HeLa cells predicts a 13.3-kDa lysine- and arginine-rich polypeptide with two extremely hydrophilic domains near the C terminus.[160] This region shows extensive homology to protamines, but there is no domain corresponding to the "RNP" consensus sequence. The D protein sequence also shows some homology to the Epstein-Barr nuclear antigen EBNA-1.[160]

A cDNA clone of the E protein was isolated by subcloning a HeLa cell cDNA library in expression vectors and selecting for translation products with an anti-Sm antibody reactive to E.[161] The E protein cDNA codes for a protein of 11 kDa.[162] Hybridization with genomic DNA from several mammalian cells indicates the presence of six to ten copies of the E protein gene per genome. The clone does not hybridize with DNA from *Xenopus* or *Drosophila*, indicating that the equivalent of E in these organisms has diverged considerably from that of mammals.[162] Preliminary analysis of the E protein gene identifies several introns and homologies to ribosomal protein genes.[163]

#### B. U1- AND U2 SNRNP-SPECIFIC PROTEINS

All three of the U1-specific proteins were cloned using expression vectors selected with the (U1)RNP antisera. A clone for the U1-specific human A protein predicts extensive sequence homology, especially at the N- and C-terminal ends with the U2-specific B" protein.[132] The two large homologous domains share 80% homology between the A and B" proteins, suggesting that these proteins were evolved after a gene duplication, and explain-

ing the presence of a class of antibodies that cross-react with these proteins.[84,138] The clone for the A protein predicts a 282-amino acid polypeptide chain with a molecular weight of just over 31,000, in close agreement with the molecular weight estimated from SDS-PAGE. The A protein shares homologous regions with other RNA-binding proteins, and has a "RNP" consensus sequence common to many such proteins.[123,170]

Clones for the U1-specific, 70-kDa protein have an open reading frame coding for a 52-kDa protein.[166] The protein is extremely hydrophilic and rich in both acidic and basic residues, and is unusually (20%) enriched in arginine. The C-terminal region of the protein is especially hydrophilic and enriched in arginine, with clusters of three or four arginines followed by a few serines or glutamic acid residues. These repeating clusters of arginines are highly homologous in nucleotide binding regions of protamines, and show homology to the "RNP" consensus sequence found in other RNA-binding proteins, suggesting that the C terminus of the 70-kDa protein binds to the U1 snRNA.[138,166,170] The 70-kDa protein will bind U1 snRNA in a total cellular RNA preparation from HeLa cells and protect a small region of U1 from nuclease digestion, suggesting a specific binding.[97] The gene for the 70-kDa protein also has a region of homology to the 30-kDa gag protein of mammalian type C retroviruses, which is an inner capsid protein that interacts with the viral RNA genome.[137,138]

Both a full-length and a partial cDNA have been cloned for the C protein.[165,166] The protein contains a region of high methionine and proline content, but it lacks a RNP consensus sequence. Southern blots suggest the gene is present in multiple copies in mammals, but as a single copy in other vertebrates.

The U2-specific A′ and B″ proteins are also cloned. As mentioned earlier, the U2-specific B″ protein has extensive homology with the U1-specific A protein.[132,165,169] Analysis of the human cDNA clone for the U2-specific B″ protein predicts a 25.5-kDa protein with three strongly hydrophilic sites that are exposed on the surface of the molecule and form the antigenic determinant.[165] One of these regions shows strong homology to a malarial circumsporozoite protein, and malaria patients frequently develop antibodies to snRNPs. The B″ protein also contains a RNP consensus sequence.[165] An isolated human clone for the U2-specific A′ protein identifies a leucine-rich and methionine-poor protein with several internal repeated runs of leucine.[168,169] The C-terminal region is hydrophilic and potentially involved in RNA binding; however, the protein lacks a "RNP" consensus sequence.[168,169]

## I. Structure of the snRNPs

### *1. Secondary Structure (Figure 1)*

The protein composition of the major snRNP particles based on the calculated protein stoichiometry suggests that the snRNPs are over 80% protein by mass. This is consistent with their density in cesium chloride or cesium sulfate at approximately 1.4 gm/cm³.[30,151] In the U1 snRNP and U2 snRNPs, approximately half the protein is from the conserved core proteins and the remainder from the snRNP-specific proteins. Although the snRNPs are only 20% or less RNA, the linear dimensions of the snRNAs are sufficient to allow interactions with many of the snRNP proteins (Figure 9). Several different approaches have been used to investigate the structure of the particles and the relationship of the snRNAs and snRNP proteins. High-resolution electron microscopy, protein and RNA cross-linking, nuclease digestion, and particle assembly on mutated snRNAs have all provided information about the structure of the snRNP particles and will be discussed below.

Sequence analysis of snRNAs predicts that each has a complex secondary structure of stem and loop structures[1] (Figure 1). Although the exact sequence of the snRNAs diverges between different species, many features of secondary structures are conserved between distant organisms, suggesting a functional importance to the structures. Although each snRNA has a stem and loop structure close to the 3′ end, the number of stems and loops varies. U1 and U2 have four stem and loop structures apiece, U4 has three, U5 has two, and U6 and U7 have only one each. By convention, the stem and loop structure nearest to the 5′ end is designated stem and loop 1, with the last stem and loop closest to the 3′ end. The consensus sequence responsible for the binding of the snRNP core proteins is in a single-stranded region in the 3′ half of the molecule. Studies described below suggest sequence elements in the stem-loop structures on the 5′ end of U1 and on the 3′ end of U2 are essential for binding of their respective snRNP-specific proteins.[89,151] Two alternative models have been suggested for the secondary structure of the U3 and U4/U6 snRNAs. The U3 models differ mostly in the details of the intramolecular base pairing in the 3′ region of the molecule (nucleotides 74 to 216), and the U4/U6 models differ in the extent of base pairing between the two molecules[1,12,13,132,133] (Figure 1C).

In the electron microscope the snRNP particles appear as approximately 10 nm in diameter.[15,16] The U1 snRNP particle visualized by high-resolution electron microscopy has a diameter of 8 nm with two protuberances about 4 nm wide. The 5′ cap is located in the main body of the particle.[16] Protein cross-linking can cross-link the B and D core proteins to the E and G core proteins, but not to each other, and the U1 snRNP-specific proteins can cross-link to each other, but not to the core proteins.[126,172] This suggests that the U1 snRNP-specific proteins and the core proteins occupy different domains and that the B and D proteins are not in contact, but rather they both interact with the E, F, and G proteins. Cross-linking of the proteins and the snRNA by UV irradiation indicates that of the snRNP core proteins, the F protein is most tightly associated with the snRNA.[126]

### *2. Nuclease Digestion Studies*

Specific protein:snRNA contacts in the snRNP particles have been investigated by nuclease digestion of the assembled snRNPs. Resistance of specific regions of the snRNA from digestion suggests that they are in tight association with protein. However,

the results can vary depending on the isolation protocols and the digestion conditions. High concentrations of $Mg^{++}$ stabilize the snRNP particles and help retain the snRNP- specific proteins in the isolated particles and protect more snRNA from digestion.[152] In general, it appears that the majority of each snRNA is associated with protein and protected from nuclease digestion. The shared core of snRNP proteins (B,B',D,D,EFG) binds to the Sm consensus sequence and protects regions around it from degradation, and the snRNP-specific proteins protect other regions of the snRNA.[39,153,154,179]

Micrococcal nuclease digestion of purified HeLa U1 snRNPs isolated under harsh conditions (CsCl-sarkosyl gradients), where only the D, E, F, and G core proteins remain bound to the snRNAs, show a protected region from nucleotides 119 to 143 toward the 3' end of the molecule.[37] This corresponds to the core binding site identified by other procedures.[18] Similar results were obtained when native U1 snRNPs, containing all the known U1-associated proteins, were digested extensively with micrococcal nuclease, suggesting that the other U1-specific proteins are bound more loosely than the core proteins. With lower concentrations of nuclease, a larger fragment of the U1 core snRNP was protected, from nucleotide 119 to 3' end of the molecule, nucleotide 165.[37] However, if a monoclonal antibody directed against the 70-kDa U1-specific protein is bound to the U1 snRNP prior to extensive micrococcal nuclease digestion, the 5' stem-loop of the U1 snRNP is resistant to digestion. These antibody protection experiments suggest that the 70 kDa protein binds to this stem-loop structure.[37] (Figure 9).

Using alternative conditions, a different pattern of RNase-resistant fragments are found in the U1 snRNP. When native U1 snRNPs, isolated under conditions where both the snRNP core and U1-specific proteins are retained, are digested with either micrococcal nuclease or RNAase III in high $Mg^{++}$ (15 *mM*), over 80% of the U1 snRNA is protected from digestion, including the 5' end.[37] Control digestions of deproteinized snRNA show digestion of almost the entire snRNA. However, in low $Mg^{++}$, only the core protein binding region is protected from digestion with micrococcal nuclease.[39,155] This suggests that the great majority of the U1 snRNA is associated with protein and that the high concentration of magnesium alters the structure of the RNP, inducing a more compact particle at high magnesium concentrations that protects the U1 snRNA from digestion.

The structure of the U2 snRNP shows a similar $Mg^{++}$-dependent conformational change. Micrococcal nuclease digestions at low concentrations of magnesium show only the core protein binding site protected from digestion, while at higher $Mg^{++}$ levels, virtually the entire 3' end of the molecule, starting with the core protein binding site, is protected.[39,155] The first 94 nucleotides from the 5' end are available for digestion under all conditions.[156] As discussed below, this correlates with assembly studies, which indicate the U2-specific proteins bind to the 3' end of the U2 snRNA.

Digestion of the U4 and U5 core particles gave similar results. In all cases, the core proteins protected a single-stranded region 15 to 35 nucleotides long, located toward or at the 3' end of the molecule, except for U2, where the protected region was located toward the center.[37] Not surprisingly, the protected region included the Sm antigen-binding sequence motif. In contrast to the other snRNPs, U6 snRNA is totally degraded during RNAase digestion of the U4/U6 snRNP particle, suggesting it is exposed on the surface of this RNP.[37]

HeLa U5 isolated by cesium chloride density centrifugation in the presence of high magnesium and separated from contaminating U1 by DEAE-Sepharose chromatography shows the core proteins D, E, F, and G, as well as a U5-specific, 25-kDa protein; however, the 100-kDa, U5-specific protein is lost in the high magnesium.[154] For unknown reasons, the B and B' do not show up on immunoprecipitates of the purified U5, though they do show on immunoblots of the same preparations. Nuclease digestion of U5 isolated under these conditions degrades approximately 80% of the molecule, and only nucleotides 78 to the 3' end, which includes the Sm binding site, are protected. This is similar to the digestion pattern seen in particles lacking the 25-kDa protein and suggests this protein covers little additional RNA.[37,153]

Nuclease digestion of U3 snRNP with RNAase A or T1 digestion of nuclear extracts, followed by immunoprecipitation with U3-specific antisera and sequencing of protected fragments, indicate that nucleotides 72 to 90, 106 to 121, 154 to 166, and 190 to 217 are protected from nuclease digestion and presumably are bound to the U3-specific proteins. Most of the protected regions are along the major stem encompassing nucleotides 74 to the 3' end of U3.[35]

### 3. Protein Binding Sites on Mutant U1 and U2 snRNAs

*In vivo* and *in vitro* assembly of altered snRNAs has helped define the sequences necessary for snRNP assembly. Sequences are eliminated from the snRNAs by either engineering of cloned snRNA genes or by site-specific nuclease digestion. Specific sequences can be degraded in both deproteinized snRNAs and in snRNP particles by hybridization with a complementary DNA probe and digestion with the RNA/DNA hybrid-specific enzyme RNAase H.

Site-specific mutagenesis of cloned *Xenopus* U1 and U2 snRNAs combined with *in vivo* assembly studies of the snRNP particles in *Xenopus* oocytes demonstrate that mutation of the core protein binding (Sm) site causes a loss of Sm immunoprecipitability, as expected, but also the loss of immunoprecipitability by antibodies against the U1-specific A or 70-kDa proteins.[39,176] However, a more sensitive competition assay between mutant and control U1 transcripts suggests a low level of A and 70-kDa protein binding to the mutant U1 snRNAs. This suggests that the core snRNP proteins stabilize the interactions of A and 70 kDa (so they could survive immunoprecipitation), but are not absolutely required for it.[178] The requirement for the binding of

the Sm core proteins (D, E, F, G, and presumably B) for stabilization of the snRNP-specific proteins is also observed with U2, where deletion of the Sm binding site prevents binding of U2-specific proteins, as measured by immunoprecipitation.[28]

The binding of the core proteins absolutely requires the Sm binding site, but the 3' stem and loop of U1 (nucleotides 143 to 158), which contains a sequence conserved in all five of the major Sm-reactive snRNAs, stabilizes the binding of the core protein complex to the Sm binding site.[28,178]

Binding of the 70-kDa and A proteins in the U1 snRNP requires the three 5' stem-loop structures, especially the stem and loop closest to the 5' end. Deletion of that loop (nucleotides 18 to 48) causes total loss of immunoprecipitability using anti-A or anti-70-kDa antibodies when assembly occurs *in vitro*, and total loss of the 70-kDa protein and greatly reduced binding of the A protein when the mutant snRNP is assembled in *Xenopus* oocytes.[178] Deletion of the other two 5' stem and loop structures reduces the binding of the A and 70-kDa proteins in the *in vivo* system, but shows little effect in the *in vitro* extract. The differences may reflect the different assembly conditions *in vivo* and *in vitro*.[178] Taken together, the data suggest that the predominant binding site for the 70-kDa and probably the A protein is located on the most 5' stem and loop structure, but the other two stem-loop structures near the 5' end and the core snRNP proteins stabilize the interaction of these two proteins with the U1 snRNP.

Details on the structure of *Xenopus* U2 were elucidated by *in vitro* mutagenesis of cloned U2 genes and subsequent injection into *Xenopus* eggs.[9,29,30] Substitution of a 12 base sequence centered around the conserved core protein binding sequence motif not only abolishes binding by the Sm antigen (and presumably the entire core protein complex), but by a U2-specific (probably A') protein as well, while substitutions or deletions in the two 3' terminal stem and loop regions abolishes binding by the U2-specific protein only.[9,30,178] This suggests that while in the U1 snRNP, the 5' stem loops are essential for binding of the snRNP-specific proteins; in the U2 snRNP, the U2-specific proteins bind to 3' terminal stem-loops.

### J. Coexistence of U4 and U6 in a Single snRNP

The immunoprecipitation of U6 snRNPs, but not deproteinized U6 snRNA, which lacks a trimethylguanosine cap, by antisera against the trimethylguanosine cap was the first suggestion U6 was complexed with other snRNAs.[143,179] This suggested an interaction between U6 and other snRNPs, but did not identify which snRNP or snRNPs U6 was interacting with. Heating of snRNA and snRNPs dissociated a single species into distinct U4 and U6 bands at about 40°C, while none of the other snRNP bands showed any change even at 50°C. This suggested that U4 and U6 are associated in a common particle, and that RNA-RNA interactions rather than RNA-protein interactions were predominantly involved in maintaining this complex, since both the native U4/U6 particle and deproteinized U4/U6 dissociated at the same temperature. On the basis of extensive sequence complementarity between U4 and U6 (Figure 1C), a model was proposed where U6 is base paired to U4, except for the stem and loop near the 3' end, the extreme 5' end, and about 35 nucleotides in the interior of the molecule.[14]

In sucrose gradients the U4/U6 snRNP particle and the U4/U6 duplex sediment faster than the other snRNPs and snRNAs, respectively. Also, U4/U6 also requires higher salt concentrations to elute off the trimethylguanosine immunoaffinity columns than do the other snRNPs.[19] Despite the potential for extensive base pairing between U4 and U6, it is possible that only limited sequences of U4 and U6 actually do base pair with each other. Psoralen cross-linking of native snRNPs, followed by RNAase T1 digestion and analysis of the fragments on two-dimensional polyacrylamide gels, identified only small sequences of U4 and U6 cross-linked to each other.[177] The sequences cross-linked in these studies are nucleotides 57 to 64 of U4 and 51 to 58 of U6, both forming single-stranded loops roughly in the center of their respective molecules. The accessibility of these sequences to psoralens argues that they are uncomplexed to protein, an observation supported by digestion studies.[87,177] However, other sequences may also be involved in base-pairing between U4 and U6, but are inaccessible to the psoralens.

Kinetic data suggest that most mature U4 is bound to U6 in the nucleus, but that there may be pools of unassociated U6 in the cytoplasm. Assembly kinetics (reviewed below) suggest that U6 binds with U4 during the cytoplasmic assembly of the U4 snRNP and that the U4/U6 snRNP returns to the nucleus. Unbound U6 remains in the cytoplasm.

Native gels of nuclear extracts have identified free U4, 15S U4/U5 particles, and 25S U4/U5/U6 complexes with the majority of the U6 in the U4/U5/U6 complex.[180,181] The regulation of the assembly and disassembly of the U4/U6 particle is unknown; however, ATP is required for the formation of a U4/U6/U5 complex, and may also regulate formation of the U4/U6 particle as well.[180,181] The U4/U5/U6 complex was not detected in sucrose gradient analyses, and this suggests that the association of U5 with the U4/U6 particle is easily disrupted. Oligonucleotide directed RNAase cleavage of the 5' end of U4 blocks formation of both the U4/U6 and the U4/U6/U5 complexes, suggesting this region is involved in the interaction.[180]

The yeast equivalents of U4 (snR 14) and U6 (snR6) are also base paired in a manner similar to that for mammalian U4 and U6. Preheating to 65°C or use of denaturing electrophoresis gels abolished this complementary base pairing. The sequences of both yeast snRNAs corresponding to the sites of psoralen cross-linking between U4 and U6 are highly homologous to that of their mammalian counterparts.[47]

## II. CYTOPLASMIC ASSEMBLY AND NUCLEAR LOCALIZATION

### A. Transcription and Association with the La Antigen

#### 1. *Transcription*

With the exception of U6, which is transcribed by RNA polymerase III, the snRNAs are all transcribed by RNA polym-

erase II.[37,49,50,80,181,183] Transcription begins with an A residue, and the newly transcribed snRNAs are capped with a 7-methylguanosine like other polymerase II transcripts.[6,26] As discussed earlier, the regulatory elements of the snRNAs differ from those of mRNA genes, and this may reflect the high rates of transcription required of these genes. Several of the snRNAs are transcribed as precursors that are several nucleotides larger than the mature nuclear snRNAs.[154] The extra nucleotides are removed and the cap is hypermethylated to a 2,2,7-trimethylguanosine during maturation and snRNP assembles in the cytoplasm as reviewed below (Figures 4 and 10).

Transcription of the snRNAs is sensitive to low levels of α-amanitin, diagnostic of transcription by RNA polymerase II.[37,39,50,182,183] Incubation of HeLa cells in 5 μg/ml α-amanitin is sufficient to inhibit U1 and U2 synthesis by more than 70%, compared to less than 10% inhibition of the RNA polymerase I or III transcription.[42,50,183] U3, U4, and U5 also exhibit α-amanitin sensitivity similar to that of pre-mRNA.[183] *In vitro* transcription of a cloned human U1 gene in a HeLa cell extract results in abnormal initiation; however, the α-amanitin sensitivity is typical of RNA polymerase II.[95] Transcription of U1 and U2 in isolated nuclei also shows a sensitivity to α-amanitin typical of RNA polymerase II.[56]

*In vitro* transcription of cloned or endogenous U6 genes display sensitivity to α-amanitin typical of RNA polymerase III.[13,14,82] As discussed above, the U6 gene displays regulatory elements typical of both RNA polymerase II and III transcription units, and *in vitro* it can be transcribed by both polymerases.[116] As expected for polymerase III genes, transcription of U6 snRNA is insensitive to α-amanitin levels (1 μg/ml) that almost completely inhibited U1 and U2 synthesis.[82,83] U6 transcription is only partially inhibited at α-amanitin concentrations as high as 50 μg/

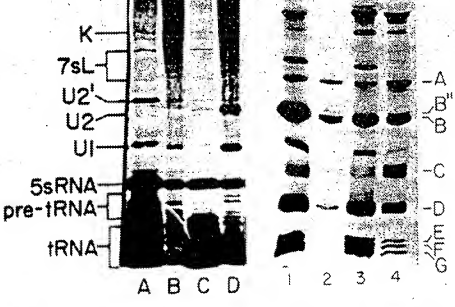


FIGURE 10. Cytoplasmic assembly and nuclear transport of snRNAs and snRNP proteins. Cells were fractionated into a cytoplasm (lanes A, C and 1, 3) and nucleus (lanes B, D and 2, 4) after a pulse label and chase. SnRNAs were labeled for 20' with $^{32}$P-uridine(lanes A, B) and chased for 45 min with 2 μg/ml actinomycin D (lanes C, D), and whole cell fractions were analyzed directly on gradient gels as described. Proteins were labeled with 50 μCi/ml $^{35}$S-methionine for 20' (lanes 1, 2) and chased with a 1000x excess of cold methionine for 90 min (lanes 3, 4). Cell fractions were immunoprecipitated with Sm antiserum and resolved on 13% gels.[197] Newly synthesized snRNAs exhibit quantitative maturation during a chase.[197] The large pools of cytoplasmic snRNP proteins show only a small fraction maturing into nuclear particles during a 90-min chase.

ml.[6] The U6 snRNA is capped at the 5' end by a methyl group.[73a] Newly transcribed U6 is transiently associated with the La antigen, and U6 can be immunoprecipitated with anti-La antibodies typical of polymerase III transcripts.[13,185] A transient association of La with U1 also occurs, despite U1 being a confirmed RNA polymerase II transcript. About 10% of the U6 in HeLa cells is immunoprecipitable with La antibodies, compared to 0.1% of the U1.[68]

### 2. Association with the La Antigen

The La antigen is a 50-kDa phosphoprotein that is associated with many RNAs transcribed by RNA polymerase III, such as nuclear 4.5S, pre-5S, pre-tRNA, and some viral RNAs. The La antigen is recognized by autoimmune sera found in a subset of SLE patients. In most cases, especially for the predominant polymerase III transcripts pre-tRNA and pre-5S RNA and U6 snRNA, the association with the La antigen is transient and restricted to the precursor forms of these RNAs. Immunoblotting experiments suggest that the La antigen is a 50-kDa phosphoprotein, and other evidence suggests that it is comprised of several isoelectric species, all of which can be phosphorylated. Isolation of the La antigen by immunoaffinity chromatography shows that it is a complex of several proteins, only one of which is reactive to La antiserum. A 64-kDa protein associated with the La antigen is required for synthesis of RNA polymerase III transcripts, but the functions and nature of the other La-associated proteins are unknown. Binding of La to the transcripts may occur on the uridine-rich 3' terminal end, although evidence from adenovirus VA RNAs suggests that the 5' end may also be involved in La binding. (For review of the La antigen and its association with polymerase III transcripts, see Reference 2.)

The ability of U6 to be immunoprecipitated by the La antigen is consistent with its transcription by RNA polymerase III transcript.[13,14] Newly transcribed U6 is immunoprecipitated *in vitro* by anti-La, but not anti-Sm antibodies.[13,14,185] About 10% of the U6 in cell extracts associates with the La antigen, and this fraction has heterogeneous 3' ends, with no "mature" 3' ends, and is undermethylated compared to mature U6. This suggests that La associates with newly transcribed U6 before many of the maturation steps, such as methylation of bases and generation of the proper 3' ends, are completed, and that the La antigen disassociates from the U6 complex before maturation is complete.[185]

Surprisingly, anti-La patient sera will immunoprecipitate a fraction of U1 snRNPs, although not other snRNPs.[186] La antisera consistently immunoprecipitates 0.1% of the U1 in HeLa, mouse, and *Xenopus* cells, but not in *Drosophila* cells, which do not have any La-reactive RNPs (showing that the La sera tested do not have (U1)RNP contaminating activity, since *Drosophila* U1 reacts with (U1)RNP antisera).[186] The La-containing U1 complexes sediment at about the same sedimentation coefficient as particles precipitated by (U1)RNP antiserum, suggesting that the La association with U1 is not due to U1 being complexed to a much larger structure that also contains the La antigen. The low percentage of U1 immunoprecipitated by La antibodies suggests that if La does associate with U1 precursors, the association is very transient. Deproteinized U1 RNA is not precipitated by the La antigen. In aqueous cytoplasmic extracts, the La antigen is associated mostly with the larger cytoplasmic precursors of U1, which are up to eight nucleotides longer than mature U1 at the 3' end, but anti-La antisera also immunoprecipitate mature-length U1 in nuclear extracts.[186]

Pulse-chase experiments suggest that the association of U1 and La in the cytoplasm is transient, because anti-La does not immunoprecipitate U1 after a chase.[185] However, La is also associated with mature-sized U1 in the nucleus, which suggests that the La antigen may play a different role than is the case for many polymerase III transcripts, where it shows a transient association with precursor RNAs, but is absent in the matured forms. A terminal C (U)$^{OH}$ at the 3' end appears to be required for La binding, and U1 (as well as U6) has such a 3' terminus, while U2 does not.[185]

## B. Cytoplasmic Assembly of snRNPs

Studies in several experimental systems demonstrate that the snRNP particles assemble in the cytoplasm, although some rapid nuclear assembly cannot be rigorously ruled out (Figure 4). Aqueous cell fractionation identifies newly synthesized snRNAs in snRNP particles in the cytoplasm within 3 min of transcription of the snRNA.[182,183,187,188] However, nuclear leakage is a problem with this type of cell fractionation (Figure 2), and rigorous studies have required techniques that prepare cytoplasmic fractions uncontaminated by nuclear components. Nonaqueous cell fractionation and enucleation of mammalian cells and manual dissection of oocytes have provided bona fide cytoplasmic fractions that confirm the presence of the newly synthesized snRNAs in the cytoplasm.[21,29,188,189]

Nonaqueous cell fractionation, which overcomes nuclear leakage by freeze-drying the cells and mechanically removing the cytoplasm, identifies precursors to U1 and U2 in the cytoplasm.[189] Cell enucleation also prepares a bona fide cytoplasmic fraction uncontaminated by nuclear material, and precursors of all six of the major snRNA species are identified (Figure 2) in the cytoplast fractions by either immunoprecipitation, hybrid selection, or Northern blotting.[188,190] Other labs, using immunoaffinity columns or immunoprecipitation with Sm or (U1)RNP antisera, confirm the presence of U1, U2, and U4 in snRNP complexes sharing determinants with mature snRNP particles in pulse labeled cytoplasmic extracts.[183,187,191] The snRNAs undergo several posttranscriptional processing events in the cytoplasm, including removal of 3' nucleotides, modification of the 5' end to form the characteristic 2,2,7-trimethylguanosine cap and methylation of bases. These processing steps and the assembly pathway of the snRNP particles will be reviewed in the following sections.

### 1. 3' Processing of Larger snRNA Precursors

U1, U2, and U4 are transcribed and transported to the cytoplasm as larger precursors that are processed in the cytoplasm to

form the mature-sized species.[92,97,152,153] The U2 precursor, U2', is a discrete species approximately 11 nucleotides larger than mature U2[28] (Figure 1D). Pulse label and chase experiments show that U2' is trimmed to mature size within 15 min although the last few nucleotides are apparently trimmed in the nucleus.[152,154,155] The maturation occurs in cytoplasm fractions, confirming that this processing occurs in the cytoplasm[155] (Figure 2).

The processing complex that trims the 3' end of U2 has been isolated on glycerol gradients of high-speed cytoplasmic supernatants.[153] A cloned U2 sequence transcribed with SP6 RNA polymerase generated a U2' precursor 11 nucleotides longer than mature U2 that was efficiently and properly processed in an *in vitro* HeLa cytoplasmic extract. However, the U2 precursor degraded with time in the whole cytoplasmic extract, but the degradative activity, a 7S complex with a generalized exonuclease activity, can be separated from the U2 processing activity in a glycerol gradient. U2' processing requires $Mg^{++}$, but not ATP, and takes about 30 min to complete. The processing activity itself appears to be a protease, but not nuclease, sensitive 15S complex with at least some specificity to snRNPs, since it does not process pre-tRNA.[153]

U1 shows a similar 3' trimming in the cytoplasm. However, the U1 precursors are a heterogeneous set of species ranging in size from one to approximately ten nucleotides larger than mature U1. Detailed studies of pulse labeled HeLa cells shows a "ladder" of U1 precursors, extending at least eight nucleotides longer than mature U1, in cytoplasmic extracts.[97,152,157] These precursors are processed to mature-sized U1 within 30 min, although some species a few nucleotides longer than mature U1 are also found in the nucleus, suggesting that the final 3' trimming may occur in the nucleus. As noted earlier, the human U1 precursors include some species with sequences that differ from the canonical U1 gene sequence, suggesting that some minor unsequenced U1 genes are also transcribed.[97] Circumstantial evidence suggests that variant 3' flanking sequences may regulate protein binding to the U1 precursors or affect assembly into snRNPs. The 3' flanking sequences of U1 all contain the Sm binding sequence $A(U)_nG$, where $n = 3 - 6$, though each variant is different.[97] It is possible that transient binding of snRNP core proteins to this region assists in the normal assembly of the particles.

U4 precursors up to seven nucleotides longer than mature U4 have also been identified in pulse labeled HeLa cell cytoplasmic extracts.[151] Like U1 and U2, they are trimmed down to mature size within about 45 min. U4 precursors have not been identified in nuclear extracts, unlike U1 and U2 precursors, suggesting that U4 3' trimming in entirely cytoplasmic. Possible U3 precursors have been identified in pulse labeled cytoplasmic extracts.[183] Larger-sized cytoplasmic precursors of U5 and U6 have not been observed. Quantitative analysis of the newly synthesized snRNA in the cytoplasm by Northern hybridization of cytoplast and karyoplast fractions with cloned probes for the U1, U2, U3, U4, and U6 snRNAs suggests they are present in exponentially growing mouse fibroblasts at a relative abundance of approximately 2% of the nuclear abundance with a half-life in the cytoplasm of approximately 20 min.[152] That is appropriate to double the number of snRNAs each cell generation for the demands of cell growth.

The extended inhibition of protein synthesis with cycloheximide interferes with the maturation of U1 and U2. The processing of U1 or U2 is unaffected by a 10-min pretreatment with cycloheximide. However, a 90-min inhibition of protein synthesis inhibits the processing of U2 by over 80% and of U1 only 20%.[153,152] This difference may reflect the fact that the pools of the U2-specific proteins are substantially less than those of the U1-specific proteins and are therefore depleted earlier. SnRNA transcription is not affected until after several hours of protein synthesis inhibition.

### 2. 5' Cap Hypermethylation

Experiments using *Xenopus* oocytes indicate that the unique 2,2,7-trimethylguanosine cap of the snRNAs is generated in the cytoplasm by additional methylations of the 7-methylguanosine 5' cap, which was added during transcription by RNA polymerase II.[28] *Xenopus* oocytes will transcribe and properly assemble snRNP, when snRNA genes are injected into the oocyte germinal vesicle or assemble snRNPs with snRNAs that are injected directly into the oocyte cytoplasm. This system has the added advantage that the oocyte cytoplasm can be manually removed to avoid problems of nuclear leakage during cell fractionation,[154,155] and has been used to great advantage for studying snRNP assembly and intracellular transport.

The injection of altered U2 genes into the oocyte nucleus indicate that upstream sequences are not required for U2 snRNP cap hypermethylation. Although deletion of the DSE or PSE abolished or substantially reduced the levels of snRNA transcription, those snRNAs that were transcribed moved into the cytoplasm and developed a typical 2,2,7 trimethylguanosine cap.[28] Deletion of U1 or U2 coding sequences, however, affected the cap trimethylation of the mutant snRNAs.[28] In particular, deletion or mutation of the Sm consensus sequence blocked the assembly of the mutant snRNAs with the common snRNP core, and the resultant snRNAs were not hypermethylated. Direct analysis of the mutant snRNAs indicated that they retained the 7-methylguanosine typical of RNA polymerase II transcripts.[28] Deletions near the 3' end of U2, which abolished binding of the U2-specific proteins A' and B'', but not binding of the Sm antigen (the core complex), had no effect on cap trimethylation. Insertion of a Sm-binding site consensus sequence, AAUUUUUGG, into two different locations in the mutant U2 gene lacking the wild-type Sm binding site resulted in both immunoprecipitability of the RNA with Sm antisera and anti-trimethylguanosine antibodies, indicating that trimethylation of the cap had occurred. The efficiency of the cap trimethylation was independent of the insertion point of the Sm binding sequence into the mutant U2 gene, although immunoprecipitation with Sm antisera was partially dependent on the location of the insertion.[28]

Like U2, cap trimethylation in U1 is dependent on binding of

the Sm proteins, since deletion of the Sm binding site caused loss of immunoprecipitability with *anti*-trimethylguanosine antibodies.[176] In contrast, deletions of the three 5' stem-loop regions and the 3' stem and loop had no effect on trimethylation of the 5' cap, indicating the important role of the Sm core protein in this process.[178]

Injection of an artificial 7-methylguanosine-capped RNA, produced by cloning of an SM binding sequence into a commercial cloning vector, into either whole or enucleated *Xenopus* oocyte cytoplasm, showed both formation of a trimethylguanosine cap and association of the Sm-reactive core proteins after 16 h of incubation in the oocyte, but not 1 min after injection.[28] This shows that cap trimethylation is dependent only on the presence of the Sm binding site and binding of the Sm core proteins (since the remainder of the RNA, except for the Sm binding site, is totally unrelated to U2), and that cap trimethylation is a cytoplasmic process, since it occurs at equal levels in both control and enucleated oocytes. Furthermore, the process takes at least a few minutes, since the RNAs could not be immunoprecipitated with either Sm or trimethylguanosine antisera 1 min after injection.[28]

The exact role of the Sm core proteins in cap trimethylation is unknown. Because the U3 has a trimethylguanosine cap, but is not immunoprecipitable by Sm antibodies, the core snRNP proteins are not absolutely essential for cap trimethylation. However, the U3 snRNP proteins may have analogous functions. Whether one or more of the snRNP proteins is the actual trimethylase, or if the proteins merely serve as a recognition site for the trimethylase, is unknown. The trimethylguanosine cap does distinguish the mature snRNAs from the pre-mRNAs and newly transcribed snRNAs in the nucleus and mature mRNA in the cytoplasm. It has been suggested that this might both prevent the snRNPs in the cytoplasm from being recognized by the translational machinery, and the mature snRNPs in the nucleus from being recognized by the transport systems that export both pre-mRNA and newly transcribed snRNAs into the cytoplasm.

### 3. Cytoplasmic Assembly of the snRNP Core Proteins

Experiments in several different systems demonstrate that the snRNP core proteins are stored in the cytoplasm in large pools of partially assembled RNA-free intermediates available for assembly with newly transcribed snRNA. Evidence from a wheat germ *in vitro* translation system suggests that the major snRNP proteins (A, B, C, D, E, F, and G) are each translated from individual poly A+ mRNAs. Sucrose gradient fractionation and subsequent immunoprecipitation of the translation products with anti-Sm sera show the majority of translation products sedimenting at about 2S, while a small fraction of assembled complexes sediment at 7S and 11S, indicating that some assembly occurred *in vitro*.[105]

Kinetic studies of mammalian cells have characterized several RNA-free assembly intermediates of the snRNP core proteins.[150,151] Sedimentation analysis of pulse-labeled cytoplasmic fractions followed by immunoprecipitation with anti-Sm sera identified the D, E, F, and G proteins in a 6S RNA-free particle[150,151] (Figure 11). Stoichiometric analysis based on isotopic labeling with specific amino acids corrected for available sequence data and quantitative staining suggests this is actually a particle of $D_2EFG$. The D' protein was also found at 20S and the B protein was in heterogeneous structures from 4S to 20S[145A,151] (Figure 11). In a pulse and chase experiment the 20S D' protein is shorter lived than the 6S $D_2EFG$ particle.[150a] Several results suggest that these are RNA-free structures and that the B protein is in structures that are independent or at best loosely associated with the $D_2EFG$ particle at 6S and the D' protein at 20S (Figures 6 and 11).

Assembled snRNP particles sediment at 12 to 16S and in the pulse-labeled cytoplasmic fractions there are no D, E, F, G proteins sedimenting in this region (Figure 11). However, after a chase, they do appear in this region, which is indicative of snRNP assembly.[150,151] Also, if cells are pretreated with actinomycin D to deplete the cytoplasm of snRNAs and halt snRNP assembly, the sedimentation distribution of the pulse-labeled snRNP core proteins is unaffected, suggesting that they are snRNA-free.[151] If the cytoplasmic extracts are immunoprecipitated with the 7I3 anti-Sm sera,[152] the $D_2EFG$ particle is precipitated at 6S and the B, but not D', protein is precipitated at 20S. This suggests that the structures containing B and D at 6S and B and D' at 20S have different epitopes and that they are in independent or at best loosely associated structures that cannot stand the rigor of immunoprecipitation.[151]

Pulse and chase kinetic experiments have also helped define the assembly order of the snRNP core proteins. In a pulse label, radioactively labeled B protein enters newly assembled cytoplasmic and mature nuclear snRNP particles before the $D'_1D_2EFG$ proteins[17,150] (Figure 10). Analysis of the assembled cytoplasmic particles focused on U1 snRNPs using a (U1)RNP antisera, so that coassembinating, but unassembled B protein in the 12-16S regions that immunoprecipitate with Sm antisera, would not interfere with the analysis. The data suggest the order of assembly is the $D'_1D_2EFG$ particle, followed by the B protein. Stoichiometric analysis suggests that two copies of the B protein add in rodent cells and in human cells one copy each of B and B'[17] (Figure 9). As discussed earlier, the B protein can be stripped from the snRNP under harsh conditions, leaving the $D'_1D_2EFG$ core associated with the snRNA, which is consistent with the independent origins of these protein complexes.[33]

Quantitative analysis has indicated that the snRNP core proteins intermediates are stored in large pools in the cytoplasm. The first indication of the large pools was obtained in *Xenopus* oocytes. Quantitative immunoblotting suggested the amount of the snRNP B protein in the oocyte cytoplasm was equivalent to the amount in 4000 somatic cell nuclei.[105,106,107] The protein stored was a 6S particle of $D_2EFG$ and B protein.[153] This was available

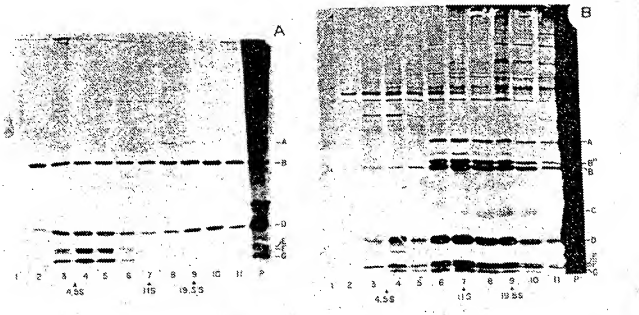


**FIGURE 11.** Sedimentation analysis of mature nuclear snRNP particles and newly synthesized snRNP proteins in the cytoplasm. Mature nuclear snRNPs were labeled with 10 μCi/ml 35S-methionine for 14 h and a nuclear fraction was analyzed on a 5 to 15% sucrose gradient. Newly synthesized cytoplasmic proteins were labeled for 60' with 30 μCi/ml 35S-methionine and analyzed on similar gradients. Gradients were collected into 11 equal fractions and a pellet and the proteins were identified by precipitation with Sm antiserum. Markers run in parallel gradients were identified by optical density and were BSA 4.3S, catalase 11S, and beta-galactosidase 19.5S. Note the snRNP particles sediment from 10 to 16S and the RNA-free proteins in the cytoplasm sediment distinctly differently. (From Sauterer, R. A. and Zieve, G. W., *J. Biol. Chem.*, in press. With permission.)

for assembly with the snRNAs, which were transcribed at a high rate following the midblastula transition. Quantitative analysis of the mammalian cells identifies a similar, though not as large, cytoplasmic pool of the snRNP proteins. Immunoblot analysis of the B protein in bona fide cytoplasmic and nuclear fractions prepared by cell enucleation indicate the B protein has a relative cytoplasmic abundance of 25% of the nuclear fraction and a half-life of approximately 2.5 h (Figure 10). This compares to the 2.5% relative abundance and 20 min half-life of the snRNAs in the cytoplasm.[106]

U1 snRNA transcribed with a SP6 RNA polymerase assembles into bona fide U1 snRNPs both in HeLa cell extracts and extracts from *Xenopus* oocytes.[20,121,138] Efficiency is greater in the oocytes, possibly due to the greater concentration of snRNP proteins. The assembly of the U1 snRNP, not surprisingly, is more efficient in a high-speed supernatant of total HeLa cell extract, which contains cytoplasmic and nuclear material, than with a nuclear extract. The assembled U1 snRNP shows the same sedimentation characteristics, $Mg^{++}$-induced compaction, nuclease resistance, and immunoprecipitability with anti-Sm and anti-(U1)RNP antibodies as the native U1 snRNP, but the reconstituted U1 is deficient in both base modification and 5' trimethylation.[20] Only 10% of the reconstituted U1 is properly capped, and approximately 10% of the U1 shows nucleotide modification to pseudouridine and methyladenosine, indicating that none of these steps are prerequisites for binding of the U1 proteins. Normal U1 assembly in the HeLa extracts is mostly complete in 15 min, and reaches maximal levels in 30 min, similar to the time newly translated U1 spends in the cytoplasm.[20]

*In vitro* studies of U1 snRNP assembly in *Xenopus* oocyte indicate the assembly process is unaffected by chelation of $Mg^{++}$, but is strongly inhibited by ATP. Salt levels as high as 500 mM NaCl do not alter binding of the core particle or A protein, but inhibits binding of the 70-kDa protein. In extracts containing the same concentration of components as found *in vivo*, assembly of U1 is essentially complete after 10 min. Binding of the Sm core proteins to the snRNAs begins immediately and increases with time.[138]

This suggests a system where newly synthesized snRNAs entering the cytoplasm are kinetically favored to assemble into snRNP particles by the large pools of unassembled snRNP core intermediates available for assembly. This is consistent with the ease at which snRNPs can be assembled *in vitro* with extracts of both mammalian cells and oocytes, and the rapid assembly

observed in mammalian cells. Kinetic analysis of the snRNP-specific proteins, however, reveals that they have significantly different kinetic behaviors than the core snRNP proteins.

### 4. Nuclear Exchange of the U1 and U2 snRNP-Specific Proteins

The high abundance of the U1 and U2 snRNPs and the availability of antisera that recognize the U1- and U2-specific proteins has allowed a kinetic analysis of the 70-kDa, A, and C U1-specific proteins and the A′ and B″ U2-specific proteins in mammalian cells. These proteins display a number of independent characteristics that are distinctly different than the kinetic behavior of the snRNP core proteins.

Quantitative immunoblot analysis of bona fide cytoplasmic fractions fails to identify cytoplasmic pools of any of the U1-specific proteins. However, during aqueous cell fractionation substantial amounts of unassembled U1-specific A and C proteins, but not the 70-kDa protein, appear in cytoplasmic fractions prepared by detergent extraction.[17,19] The data suggest that there are large pools of unassembled A and C protein in the nucleus that leak from isolated nuclei. Pulse and chase experiments done in the presence of actinomycin D to inhibit snRNA synthesis and *de novo* snRNA assembly indicate the newly synthesized A and C proteins enter the nucleus and assemble with mature stable U1 snRNP particles in the nucleus.[17] Stoichiometric analysis suggests the 70-kDa U1-specific protein is present in some, but not all, of the mature nuclear U1 snRNPs, and there are two copies of A and C in each U1 snRNP. This suggests the A and C proteins that rapidly enter the mature U1 snRNP particles in the nucleus must exchange with or replace other copies of the A and C proteins.

The pools of unassembled A protein and C proteins that leak into the cytoplasm during aqueous cell fractionation sediment in 18S to 22S and 4S to 8S particles, respectively[151] (Figure 11). Pulse-chase studies also show posttranslational modification of the C protein. During a chase there is a shift in electrophoretic mobility of the C protein corresponding to an increase of 1000 Da in molecular weight.[17,150,151]

The U2-specific A′ and B″ proteins also display independent behaviors. Cell fractionation identifies a pool of unassembled A′ protein in the nucleus, but no appreciable unassembled pool of the B″ protein. The B″ protein apparently assembles with U2 snRNP particles immediately after translation. If U2 snRNA synthesis is inhibited, this protein does not enter the nucleus. This suggests the B″ protein assembles with the U2 snRNP in the cytoplasm and enters the nucleus with the particle where it is stably associated.[19]

## C. Transport of snRNPs Into the Nucleus

After assembly in the cytoplasm, the snRNP particles return to the interphase nucleus. In both *Xenopus* oocytes and mammalian cells, large pools of unassembled snRNP proteins exist in the cytoplasm and move into the nucleus only after assembly with snRNA.[100,198] Several lines of evidence suggest that the core proteins are responsible for generating the nuclear localization signal. With limited exceptions, the major snRNAs appear transiently in the cytoplasm, where they assemble with the snRNP core proteins before returning permanently to the interphase nucleus. If the Sm consensus sequence, responsible for directing snRNP core assembly, is cloned into a heterologous RNA, it will direct snRNP core assembly and the particle will move into the nucleus. However, if the consensus sequence is mutated so that it does not assemble the snRNP core proteins, the RNA stays in the cytoplasm.[96,178] Also, removal of sequences responsible for binding of the specific proteins to the U2 snRNP does not prevent snRNP core assembly or nuclear accumulation of the U2 snRNP particle.[19] This suggests that the signals for nuclear localization are generated by the protein-protein or protein-RNA contacts between the snRNP core proteins and the RNA.[198] The nucleolar U3, which does not share the common core of snRNP proteins, but has its own unique set of proteins, must have a nuclear, and, more specifically, a nucleolar transport signal.

The secondary structure of the snRNA also affects the nuclear transport of snRNAs.[198] U2 mutants with deletions and substitutions in the stem and loop nearest the 3′ end initially had 3′ extensions on the transcripts, which were slowly processed down to the proper 3′ end over a period of several hours. Both the 3′ extended and processed transcripts could be immunoprecipitated with anti-Sm antisera, but only the processed transcripts migrated to the nucleus when purified from *Xenopus* oocytes and reinjected. The 3′ extended transcripts of these mutants were found only in the cytoplasm. Secondary structural analysis of the 3′ extended mutants predicted an interaction between the extension and the stem and loop closest to the 5′ end, causing a radical conformational change.[198] Although this change blocked the nuclear transport signal, it did not inhibit core protein binding, indicating that core protein binding alone is not the only factor involved in nuclear transport of snRNPs.

Transport of newly assembled snRNPs into the nucleus after maturation shares many features with the nuclear transport of large proteins. The snRNP particles are too large to enter the nucleus passively, so they must be actively transported across the nuclear pore.[199,200] Studies in a variety of experimental systems demonstrate that the binding to the nuclear pore and uptake into the nucleus is a two-step process and is a result of a specific amino acid sequence.[201,202] The karyophilic protein or particle first binds to the nuclear pore and then is actively transported into the nucleus. The karyophilic signals studied do not show absolute sequence conservation, but do have some common features: a short region of basic amino acids, usually flanked on one or both sides with a few hydrophobic residues.[203,204] In this sense, they are similar to signal sequences in secreted proteins, which do not show sequence homogeneity, but do show similarities in the arrangement of hydrophobic and polar amino acids within the sequence. The experimental data suggest that snRNP core assembly generates such a karyophilic signal, which then binds to the nuclear pore and triggers nuclear uptake.[198]

Cell fractionation data suggest that snRNP core assembly in

the cytoplasm occurs in a soluble compartment, and that targeting of the snRNP particle to the pore is a diffusion-mediated step, rather than an active, cytoskeletal-mediated process.[93] Pretreatment of cells with a wide variety of inhibitors of intermediary metabolism or of the cytoskeleton all failed to affect snRNP maturation or transport into the nucleus. Only hypertonic medium, prolonged inhibition (60 min or more) of protein synthesis, or cold shock blocked maturation and transport. Treatment of cells with medium adjusted to approximately twice the normal osmolarity with either salts or sugars (the addition of 180 mM NaCl or 360 mM sorbitol to normal medium) blocked maturation and transport of snRNPs into the nucleus (Figure 12). This effect is completely reversible when the cells are reincubated in normal medium. Hyperosmotic medium withdraws water from the cytoplasm and the entire cell shrinks in volume. One hypothesis is that the collapse of the cytoplasmic matrix under these conditions blocks the normal diffusion of the particles in the cytoplasm.[92] This treatment also induces foci of Sm staining in the cytoplasm, which may be aggregates of the snRNP core proteins (Figure 12F). The inhibitory effect of extended protein synthesis inhibition on snRNP maturation is first seen with the U2 snRNP. This effect is likely the result of the depletion of essential proteins and may suggest a role for the U2-specific proteins in the maturation of this particle. The inhibitory effects of cold shock are probably due to a general arrest of cell metabolism.[92]

### D. Cellular Localization of snRNPs

Cell fractionation studies of the major snRNPs originally

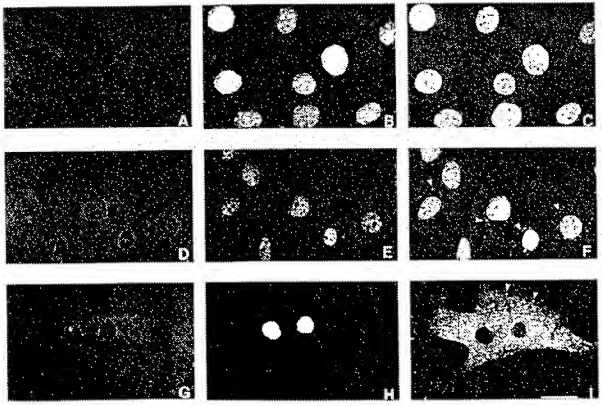


FIGURE 12. Hyperosmotic medium inhibits the intracellular transport of snRNPs during interphase and mitosis. Nil 8 human fibroblasts were exposed to hyperosmotic medium by the addition of 180 mM NaCl to the culture medium. Control and experimental cells were examined by phase microscopy (A, D, G), Hoechst staining (B, E, H), and indirect immunofluorescent staining with the Y12 anti-Sm monoclonal antiserum I h after the alteration in the tonicity of the culture medium. Control interphase cells (A, B, C) show typical nuclear localization of the Sm antigen. Interphase cells exposed to hyperosmotic medium (D, E, F) show the development of regions of Sm staining in the cytoplasm (arrows in F). Hyperosmotic medium prevents the action of the cleavage furrow and the reformation of the daughter nuclei in dividing cells (G, H, I). The chromatin remains condensed and the mature nuclear snRNPs stay dispersed throughout the cytoplasm as the cell spreads back out on its substratum. (Bar, 10 μm.)

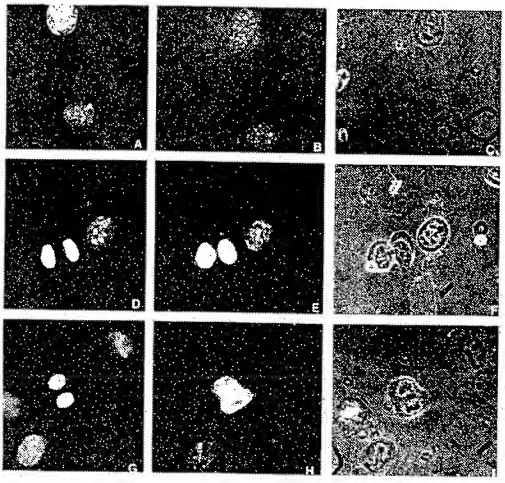

FIGURE 13. Indirect immunofluorescent staining of snRNP particles during interphase and mitosis in NIH hamster fibroblasts. Hoechst (A, D, G), indirect immunofluorescent staining (B, E, H), and phase micrograph (C, F, I) of interphase NIH hamster fibroblasts stained with (U1)RNP antiserum (A, B, C), and of interphase and late mitosis cells with the Y12 anti-Sm monoclonal antibody (D, E, F) and an anaphase cell with anti-Sm monoclonal antibody (G, H, I).

identified the five major snRNPs, U1, U2, U4, U5, and U6 in the nucleoplasm, and U3 confined to the nucleolus.[34] The availability of specific antisera have allowed more detailed localizations of the specific snRNP particles. In the interphase cell, the vast majority of snRNPs observed by immunofluorescence reside in the nucleus[207-210] (Figures 12 and 13). Although Western blotting identifies a pool of snRNP core proteins in the cytoplasm, they are difficult to detect by immunofluorescent staining, possibly because of their low concentration.[206,211] In the interphase nucleus, computer-aided image analysis of immunofluorescent images indicates that Sm-reactive snRNPs colocalize with (U1)RNP-reactive (U1) snRNPs into domains that appear as speckles and are often described as a punctate staining pattern.[208] Although other nuclear antigens also localized to distinct domains, there was only partial overlap with Sm and La antisera, and virtually no overlap between Sm and an anticentromere antigen. There was also considerable overlap between Sm staining and an antiserum that recognized a 107-kDa nuclear matrix protein, though the two antigens showed radically different distributions during mitosis.[105]

Immunoelectron microscopy of snRNPs (Figure 14) showed them clustered into nonchromatin regions in interphase nuclei.

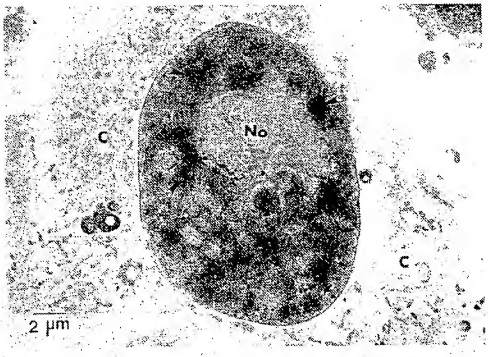


FIGURE 14. Immunoperoxidase staining of snRNPs in CHO cells: 0.5-μm section of a CHO 400 cell showing the distribution of snRNP (anti-Sm staining) via immunoperoxidase staining. Arrowheads point to nuclear regions enriched in snRNPs. The nucleolus (No) and cytoplasm (C) show no stain. This section has not been counterstained. (Micrograph courtesy of D. L. Spector, Cold Spring Harbor Laboratory.)

but also existing in smaller amounts in the general nucleoplasm, suggesting that the snRNP clusters are interconnected to form a reticulum in the nucleoplasm.[208,210,213] Some, but not all, clusters of snRNP particles in the nucleus colocalize with the interchromatin granules in the nucleoplasm.[35] The EDTA bleaching procedure highlights the interchromatin granules, and previous studies suggested they are stable RNP structures that do not contain newly transcribed pre-mRNA.[214] They may represent either storage forms of the snRNPs or processing scaffolds for the pre-mRNA.[35] The *myc* oncogene product colocalizes in the same reticulum.[212] In chromatin spreads, both Sm and RNP antigens appear on RNP fibrils being transcribed from the DNA.[211,215] There was no such association with ribosomal genes, however.

Immunofluorescence studies of *Drosophila* salivary gland polytene chromosomes using anti-U1 and -U2 sera show both localized to chromosome puffs, in a degree dependent on transcriptional activity as measured by uridine incorporation and treatment with agents that stimulate or repress transcription at specific loci. All transcriptionally active bands appear to label with U1 and U2 antisera, and inactive chromosome bands show no U1 or U2 down to the level of detection with immunofluorescence.[216]

Cell fractionation data have indicated that the snRNP particles are enriched in an insoluble nuclear matrix fraction. When nuclei are extracted under conditions that remove the chromatin, a fibrillar internal network remains, which is highly enriched for the snRNP particles.[217,218] The 70-kDa U1-specific protein is also enriched in nuclear matrix preparations. If the U1 particles are selectively eluted from isolated nuclei by incubation at elevated temperature and pH, the 70-kDa protein remains in the nucleus attached to this matrix.[219,220]

A small fraction of the snRNAs are found covalently linked to chromosomal DNA through a 3′ to 5′, DNA-RNA linkage. These snRNAs appear to be in a dynamic equilibrium with the more abundant snRNAs in the nucleoplasm. The significance of the snRNA-DNA linkage is not known, although it was suggested that it modulates the tertiary structure of the chromatin.[221]

## E. SnRNPs During Mitosis and Meiosis

In most higher eukaryotes the nuclear envelope breaks down at the onset of mitosis. When the chromosomes are fully condensed near the end of prophase, the nuclear lamina disperses and distributes throughout the cytoplasm (Figure 13H). It remains dispersed until late telophase, when it begins reforming on the

surface of the decondensing chromosomes of the daughter nuclei (Figure 13E). When the envelope breaks down, the stable nuclear snRNPs distribute throughout the cytoplasm, with only a small fraction remaining associated with the surface of the chromosomes.[219,222-226] The punctate distribution of the particles observed in the interphase nucleus is lost, and the particles appear uniformly distributed throughout the cytoplasm.[222-224] The snRNP particles retain their normal antigenicity and protein composition during this time as determined by immunoprecipitation and indirect immunofluorescent staining with the SLE autoimmune antisera.[222,223] The 70-kDa protein associated with the U1 snRNP and thought to bind to the nuclear matrix in interphase cells remains associated with the dispersed U1 snRNPs during mitosis.[219] Sedimentation analysis indicates that the snRNPs remain in heterodisperse structures ranging from the individual 12S particles up to structures of over 100S. Selective extraction of metaphase cells with cytoskeletally stabilizing buffers suggests that approximately 40% of the particles are soluble in the cytoplasm, but that the remainder are associated with large insoluble structures.[227]

The snRNP particles begin returning to the daughter nuclei immediately after the chromatin begins decondensing in telophase, and the particles return quantitatively to the daughter nuclei during early G1 (Figure 13). The rapidity with which the particles begin clustering in the region of decondensing chromatin suggests an intrinsic affinity for the nuclear environment, and suggests that it is not necessary for all the mature snRNP particles to be transported through the nuclear pores to enter the daughter nuclei.[222,223] However, some snRNP particles remain in the cytoplasm until the nucleus is nearly fully assembled, and it is likely that they could be transported into the nucleus through the nuclear pores.[222]

In an effort to analyze the return of the mature particles to the nuclei following the completion of mitosis, Zieve and Sliziky[222] investigated a variety of metabolic inhibitors for their ability to interfere with this movement. None of the well-characterized reagents that disrupt the cytoskeleton or inhibit cellular metabolism blocked this movement. Only the exposure of cells to hypertonic medium inhibited the return of the snRNP particles to the daughter nuclei. When cells in anaphase or telophase were exposed to hypertonic medium, the further movement of the chromosomes and the activity of the cleavage furrow were inhibited. However, the cells flattened out as if returning to interphase. The chromatin remained condensed and the stable snRNP particles remained dispersed throughout the cytoplasm (Figure 12I). This suggests that the return of the particles to the daughter nuclei is the result of diffusion in the cytoplasm coupled with specific binding sites in the chromatin and at the nuclear pores.[222] As discussed earlier, hypertonic medium also blocks the movement of the newly synthesized snRNP particles into the interphase nucleus[192] (Figure 12F). This is consistent with the hypothesis that the return of the newly synthesized snRNP particles during interphase and the mature stable particles at the completion of mitosis occur by similar mechanisms.

Sea urchin eggs exhibit the only recorded situation where assembled snRNP particles remain in the cytoplasm.[223] *In situ* hybridization of U1 antisense RNA to thin sections of sea urchin oocytes, eggs, and embryos found most, but not all, of the U1 snRNA in the nucleus during oogenesis. However, after germinal vesicle breakdown, the maternal snRNA remained in the cytoplasm and stayed there throughout the early stages of development. New snRNP particles were assembled from U1 transcribed from genes in the embryo and displayed a normal maturation cycle, while the maternal snRNPs remained in the cytoplasm. This suggests that the maternal snRNPs are altered in some way so that their nuclear localization signal is destroyed and that only newly synthesized snRNP particles accumulate in the embryonic nuclei. Psoralen cross-linking indicated that some of the U1 snRNPs in the cytoplasm of the oocytes and early embryos were base paired with poly A+ RNA in the cytoplasm.[226] Immunoprecipitation and hybridization analysis of this poly A+ RNA indicated that it is the unusual sea urchin transcripts that contain single-copy DNA sequences interspersed with repetitive sequences that are found in the oocyte cytoplasm.

## III. FUNCTIONS of snRNPs

In the cell nucleus, the snRNPs function in the processing of newly transcribed RNA. The major snRNPs participate in the removal of introns from premessenger RNA by RNA splicing (U1, U2, U4/U6, and U5), the 3' end processing of nonadenylated (U7) and adenylated (U11) pre-mRNAs and preribosomal RNA processing (U3, U8) in the nucleolus. The roles of the snRNPs in these functions are reviewed below.

### A. Pre-mRNA Splicing in the Nucleoplasm

The basic features of snRNP-mediated pre-mRNA splicing have been reviewed extensively in recent literature.[8,13,16,227,228] This is one of three types of RNA splicing that are now described.[229] All the splicing reactions involve a two-step reaction with an initial cleavage at the 5' splice site followed by a cleavage at the 3' site where both cleavage reactions are transesterification reactions with the cleavage at one site coupled to a ligation at another site. Also, conserved sequences at the splice junctions are critical for the specificity of the reactions[229] (Figure 15). The pre-mRNA splicing in the nucleus is unique because of the requirement for snRNPs in the splicing reaction. However, a reaction that is similar in many details occurs in organelles without the participation of snRNPs, and this suggests that in specific circumstances the functions provided by the snRNPs in the nucleus can be provided by intramolecular interactions of the RNA of the substrate itself.[229] In the following review we focus on the roles of the snRNPs in the assembly and function of the splicing complex in the cell nucleus, the spliceosome, which includes U1, U2, U4/U6, and U5 snRNPs.

#### 1. *Substrate Requirements for* In Vitro *Splicing*

The analysis of pre-mRNA splicing has relied heavily on *in*

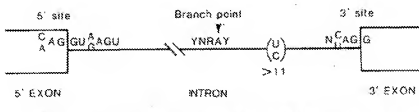


FIGURE 15. Cartoon model of conserved sequence motifs in the mammalian intron. The conserved 3' and 5' splice sites and the branch point sequences are illustrated. In mammalian cells, but not yeast, there is also a necessary polypyrimidine tract between the branch point and 3' splice site. R is a purine, Y is a pyrimidine and N is any nucleotide.

*vitro* systems that accurately splice an added substrate.[31,37,227,228] Typically, the substrate used in an *in vitro* splicing reaction contains a truncated, cloned, pre-mRNA transcript, such as a β-globin transcript or adenovirus late transcript containing the first intron and the two flanking exons. However, other transcripts are also used as splicing substrates.[229-233] When the splicing substrates are incubated in a nuclear extract (normally from HeLa cells) in the presence of $Mg^{++}$, monovalent cations, and ATP, the substrate premessenger RNA is spliced accurately.

The splicing reaction in mammalian cell extracts can be divided into three stages. After an initial lag period of as long as 45 min in the *in vitro* reaction, the splicing intermediates appear and spliced mRNA accumulates with linear kinetics for the next 2 or 3 h.[120,231] ATP is required for some events during the initial lag period when the splicing apparatus, the spliceosome, is assembling.[230-235] The next stage in splicing results in cleavage of the precursor RNA at the 5' splice site and the ligation of the 5' end of the intron to an adenosine residue within the intron, the branch point, by a 5' to 2' linkage to form a structure described as a lariat. This results in the cleavage of the substrate into two pieces, the 5' exon and a RNA containing the entire intron and the 3' exon.[230,235,236] Finally, the 3' splice site is cleaved and ligated to the 5' splice site of the first exon and the intact intron, in a branched configuration, and spliced exons accumulate as the previous intermediates decline in quantity. Typically, in an *in vitro* splicing reaction, appearance of the 5' exon and the intron-3' exon complex occurs after 30 min, and the spliced exons and intact introns appear about 15 to 30 min later.[230,233,236]

The lariat structure of the excised intron, with the 5' nucleotide of the intron linked to an adenosine residue near the 3' end of the intron by a 5' to 2' linkage, is characteristic of the type II splicing reaction. It also appears in the self-splicing introns found in organelles.[229] This unusual structure was initially identified because it migrates abnormally slow on electrophoresis gels, it blocks reverse transcription in primer extension experiments at the same site near the 3' end of the intron, and is resistant at that site to digestion by RNAase P1.[230,236] Analysis with site-specific RNAase H cleavage using synthetic oligonucleotides complementary to regions in the intron indicated that an adenine near the 3' end of the intron had a 2' to 5' branch with the 5' end of the intron.[230,233,236] Excised introns have been identified *in vivo* in both a linear and a lariat configuration, suggesting that a debranching activity is present in the nucleus.[233,238]

Three sequences play vital roles in splicing: the 5' splice junction (the semiconserved consensus sequence is AG↓GUAAGU, with the GU essentially invariant); the 3' splice junction (consensus sequence is CAG↓G, with an invariant AG); and the branch point near the 3' end of the intron[34,228] (Figure 15). The branch point in higher vertebrates forms only a weak consensus of PyNPyPu(A)Py, with Py representing a pyrimidine, Pu a purine, and N any nucleotide, and the (A) being the actual site of the 2' to 5' phosphodiester bond forming the branch, while in yeast the branch point sequence is an absolutely conserved sequence of UACUA(A)C, with the branch point at the (A). The branch point usually occurs about 30 nucleotides upstream from the 3' splice site, and introns appear to be a minimum of 60 nucleotides in length.[236] In mammals there is an essential polypyrimidine stretch upstream of the 3' splice site; between the 3' site and the branch point, that is not necessary in yeast.[235,236] Mutants of essential sequence motifs in yeast usually inactivate the introns; however, in mammalian cells it often activates a nearby cryptic site.[9,32,43,245-247] Introns vary enormously in size and can be up to several kilobases in length, while exons tend to be no more than several hundred nucleotides long. Estimates suggest there are approximately ten introns per average pre-mRNA, of which 10,000 exist at a given time in the HeLa cell nucleus, for a ratio of U1 and U2 per intron of about 5:1.

High-resolution negative stain has visualized the splicing complex, the spliceosome, both in Miller spreads of transcription complexes and of spliceosomes isolated from sucrose gradients.[235,247] In negatively stained transcription complexes prepared from *Drosophila* chromosomes, a 10-nm particle is occasionally seen at the 5' splice site even before the 3' site is synthesized. When the 3' splice site is transcribed it associates with a 25 nm particle, and the 5' and 3' sites quickly coalesce to form a 40-nm particle as the intron is looped out.[235] Spliceosomes isolated from mammalian *in vitro* splicing extracts appear as particles 40 to 60 nm in diameter with several subunits that may be the individual snRNPs. Spliceosomes in yeast cell extracts are similar to those in mammalian, although they appear to be somewhat smaller.[240]

### 2. Spliceosome Assembly and Composition

Even before the discovery of pre-mRNA splicing, the local-

ization of snRNP particles in isolated hnRNP fractions and the ability to cross-link snRNAs to hnRNA with psoralen suggested that snRNAs were involved in the maturation of pre-mRNA.[200,260,270,283] With the discovery of splicing, several lines of evidence initially suggested that the snRNPs played a vital role in splicing. After the conserved sequences at 5′ splice sites were identified, sequence complementarity between regions near the 5′ end of the U1 snRNA and the consensus sequence were quickly noted and a role for U1 snRNA in pre-mRNA splicing was proposed.[283,284] Although the original models for the splicing reaction were incorrect, other data described below confirmed a vital role for the U1 snRNP in recognition of the 5′ site through RNA:RNA base pairing. Also, initial analyses of the splicing reactions in *Xenopus* oocytes indicated that Sm antisera directed against the common core of snRNP proteins inhibited the splicing reaction, suggesting the snRNP particles played a vital role in splicing.[285] Recent studies have used a variety of sophisticated approaches to investigate the essential roles of the major nucleoplasmic snRNPs in the splicing reaction (Figure 5).

Analysis of splicing reactions by sedimentation in glycerol gradients or electrophoresis in nondenaturing acrylamide gels reveals that the splicing complex, the spliceosome, is a dynamic multicomponent complex of about 40S in yeast and 50 to 60S in mammalian systems. The functional complex contains the pre-messenger RNA and up to five of the snRNP particles[61,62,115,180,233-235,269] (Figure 5). In some studies the complexes are treated with either alkaline pH or heparin to reduce nonspecific absorption of cellular proteins, and this often removes U1 from the splicing complex and reduces the size of the complex compared with untreated controls.[108] Recently, spliceosomes have been purified by affinity chromatography of derivatized substrates and shown to contain all five of the nucleoplasmic snRNPs.[62,226-230]

Kinetic experiments with *in vitro* splicing extracts reveal discrete steps in the assembly of mammalian spliceosomes which help identify the roles of the individual snRNPs in the splicing reaction.[145,268,232-259] Upon addition of a pre-mRNA to the splicing extract, a 22S complex forms almost immediately. This complex forms in ATP-depleted extracts, and detailed analysis of this complex shows the presence of the U1 snRNP and RNA binding proteins, including the hnRNP proteins[60,259,260] (Figure 5). However, when complexes are treated with conditions to reduce nonspecific absorption before analysis, the U1 snRNP is lost from the complex and migrates slightly faster in nondenaturing gels and on gradients.[145,190,241] U1 association occurs immediately after addition of substrate RNA to the splicing extract and reaches a peak within 1 min, and the formation of this complex is inhibited by ATP.[257] RNAase T1 digestion of this complex shows a 15-nucleotide protected region encompassing the 5′ splice junction.[60] Although the 5′ splice site is required for optimal association of U1, a limited amount of the U1 snRNP can associate on substrates lacking a 5′ splice site provided the 3′ region is intact, indicating that interactions with other regions of the substrate or other snRNPs are partially responsible for stabilizing the U1 binding.[62,259] If the U1 snRNP is depleted from the extract, or if requirements for U1 snRNP binding are totally eliminated, the formation of the 22S complex is abolished.[62,258,259]

Within 5 min, a 35S complex forms as the 22S complex disappears. Formation of the 35S complex is ATP-dependent and requires the U2 snRNP and at least two other factors[233-235] (Figure 5). Removal of the 5′ region of the U2 snRNA by RNAase H cleavage, or deletion of either the polypyrimidine tract near the 3′ splice site or of the branch point sequence, blocks formation of the 35S complex.[145,233,262,266] The 35S complex contains U1 and U2. However, on native gels, where extracts are treated with either heparin sulfate or high pH, the ATP-dependent complex containing U2 sedimenting at 22S and lacks U1 snRNP.[145,233,258,264] The U2 snRNP does not bind to the 22S splicing complex alone, but requires a U2 association factor, U2AF, and possibly another factor is also required for 3′ splice site recognition and U2 snRNP binding.[232-234,253,266]

By 15 min, in *in vitro* splicing extracts, the 35S complex starts to disappear and is replaced by a 50S to 60S complex containing all five of the nucleoplasmic snRNPs, while the first splicing intermediates are formed[145] (Figure 5). In heparin or high pH, this complex sediments between 30 to 35S.[140] Formation of this complex is blocked if nucleotides 29 to 42 of U2 (the region that may base pair to the branch point) or the second loop (nucleotides 65 to 84) of U4 are degraded by RNAase H and complementary oligonucleotides, although the 35S prespliceing complex can still form.[145,262,262,266] A variety of kinetic studies suggest that the addition of U4, U6, and U5 appears to be a single-step process involving the association of a 25S particle containing the three snRNPs, although evidence also exists for an early, possibly transient association of U5 with the 3′ splice site, where it is involved in assisting the binding of U2 to the branch site.[130,180,259,266] U4 and U6 appear predominantly in the U4/U6/U5 particle, with some in a 15S U4/U6 complex.[130]

Both the U4/U5/U6 and U4/U6 complexes appear in the absence of pre-mRNA substrate, indicating that their formation is independent of the formation of spliceosome precursors.[18,19,180,266] The presence of ATP increases the level of the 25S U4/U5/U6 complex relative to the U4/U6 particle or free U4, U5, and U6. Since ATP is required for the binding of U6, U5, and U6 to the spliceosome, it is possible that the ATP is required to assemble the U4/U5/U6 complex, rather than the actual association of the complex with the forming spliceosome, or ATP could be required in both steps.[182] Oligonucleotide-directed RNAase H cleavage of U4 prevents formation of both the U4/U6 particle and the 25S U4/U6/U5 complex, as well as strongly inhibiting spliceosome formation and splicing, indicating the crucial role of U4 interactions with other snRNPs in splicing.[182] The 50S complex containing all five snRNPs is the mature spliceosome, and it remains until the splicing reaction is complete. Bound to this complex are some of the precursor RNA, the 5′ exon, the lariat, and most of the product intron and spliced exons.

The 50S mature spliceosome complex can still form if the 3' AG at the 3' splice junction is deleted (as long as the polypyrimidine tract is present), but it is extremely slow and inefficient.[254,255] Addition of just a few nucleotides in the 3' exon increases efficiency of the 50S spliceosome formation to near normal levels. In both cases, the 5' exon is cleaved and the lariat intermediate is formed, but exon ligation does not occur until there are 12 nucleotides or more in the 3' exon (the mutant RNA with the 3' AG deleted has no 3' exon at all), presumably because of required sequences or structure in the 3' exon.[255] If the 5' splice site is deleted, some U1 will associate with the substrate and a 35S complex can be detected; however, the 50S spliceosome does not assemble and there are no cleavage reactions.[250,260,268] Thus, formation of the complete spliceosome and splicing requires intact 5' splice sequences, the branch point (or a cryptic one), and the polypyrimidine tract near the 3' end. The presence of a 3' splice junction is not absolutely required for spliceosome formation, provided the polypyrimidine tract is intact,[260] but it increases the assembly of the mature complex, and it, as well as some nucleotides in the 3' exon, are required for complete splicing.

The final step in splicing involves release of the excised intron in a lariat form in a 29S complex on native gels that contains U2, U5, and U6.[152,181] Surprisingly, U4 is not found in the complex, indicating that it dissociates from U6 at some point prior to intron excision and is recycled to form new U4/U6 and U4/U6/U5 complexes later.[180,181,252,258] However, recent protocols using affinity purification of the U4/U6 snRNP particle with complementary biotinylated nucleotides suggests that the U4 snRNP particle is not lost from the spliceosome during the splicing reaction.[256] The spliced exons appear in heterogeneous forms, probably associated with protein.

Interactions between various snRNPs are vitally important in spliceosome formation and integrity. Nuclease-protected fragments from the 5' splice site can be immunoprecipitated from splicing reactions using either antibodies to U1 or U2, and 3' splice site regions can be immunoprecipitated with anti-(U1)RNP antibodies, indicating an interaction between U1 and U2.[145,274] The size of the protected fragments are larger in the 35S complexes or mature spliceosomes than in 22S complexes containing only the U1 snRNP, suggesting a rearrangement of the complex as it assembles.[145,260] Likewise, depletion of a splicing extract of U1 or deletion of the 5' splice site inhibits the binding of U2 and U4/U5/U6.[254,259] Similarly, RNAase H cleavage of U2, U4, and U6, singly or in concert, or deletion of the branch point and 3' splice site, while allowing U1 binding to the 5' splice site, prevents spliceosome assembly and the 5' splice site remains susceptible to degradation by RNAase T1.[254,262] Thus, there are strong interactions between U1 and U2, and U4, U5, and U6 also interact with each other and stabilize the entire spliceosome. The importance of snRNP-snRNP interactions is further demonstrated by formation of a U2/U4/U5/U6 particle in splicing extracts at high salt, even in the absence of substrate RNA, to form a particle called the pseudospliceosome.[268] As mentioned earlier, high salt causes the unassembled pools of cytoplasmic snRNP core proteins to coalesce into foci in the cytoplasm, suggesting high salt promotes specific homotypic interactions between the snRNP proteins.[192]

Spliceosome assembly and splicing in yeast is similar to that in mammalian systems. Assembly of yeast spliceosomes requires intact 5' and branch point sequences, as in mammals, although yeast intron junctions lack the polypyrimidine tract near the 3' end that is found in higher eukaryotes, and the consensus 5' splice site and the branch-point sequence are more highly conserved than in mammals.[243,270] The mature yeast spliceosome sediments at 40S instead of 50 to 60S for mammalian spliceosomes.[148] The assembly of yeast spliceosomes is similar in kinetics and snRNP requirements to that in mammalian systems.[41,42] Within 30 s of addition of substrate to yeast-splicing extracts, a complex called band III forms. In native gels, this complex contains only yU2, but affinity purification procedures and a recent modification of the native gel procedure that use less stringent conditions also show the presence of yU1.[5,22,253b] Kinetic studies suggest that the yU1 binds to the substrate before the yU2 and that sequences at both the 5' splice site and at the branch point are necessary for yU1 binding.[63] Band III, therefore, corresponds to the 35S spliceosome precursor in mammalian systems.

Band II is a precursor to a complex called band I, which is the functional 40S spliceosome. Like its mammalian counterparts, band I contains the full complement of yeast spliceosomal snRNPs yU1, yU2, yU4/yU6, and yU5.[217,238] The full set of snRNPs and an intact branch point are required for band I formation.[257,258] Band I contains both splicing intermediates and products. Similar to the situation in mammalian cells, an ATP-dependent yU4/yU6/yU5 complex forms and can be immunoprecipitated with an antisera to the yU5-specific, 260-kDa RNA 8 protein.[156] In the final stage of the splicing reaction the yU4 is released and a band II complex containing the majority of the splicing products and yU1, yU2, yU5.[257,258]

We now review the data on the requirements for the individual snRNPs in the assembly in the spliceosome.

### 3. Roles of snRNPs in Splicing

#### A. U1

Substantial data have accumulated indicating that the U1 snRNA base pairs with conserved sequences at the 5' splice site during the initial stages of spliceosome assembly, and there are growing indications that the U1 snRNP also interacts with sequences near the branch point as well. This suggests that binding of the U1 snRNP to the pre-mRNA substrate is a critical first step for aligning the 5' and 3' splice sites.[82] Base pairing of the U1 snRNA to the pre-mRNA was first detected by psoralen cross-linking.[201,221] Additional evidence for a vital role of U1 in splicing was initially obtained in *in vitro* splicing systems when it was found that anti-Sm and anti-(U1)RNP sera inhibited splicing when added to splicing systems at levels sufficient to quantitatively remove snRNPs.[203-205] Surprisingly, anti-U2 se-

rum had no effect, suggesting a different arrangement of the U1- and U2-specific proteins in the spliceosome. Addition of anti-(U1)RNP antibodies to an *in vitro* splicing system prevented the formation of the 60S spliceosome, and both anti-(U1)RNP and anti-Sm antibodies immunoprecipitated splicing intermediates from the 60S spliceosome, indicating that U1 is a component of the spliceosome and is necessary for its assembly.[220,260,277]

Although sequences near the 5′ end of U1 are complementary to both 3′ and 5′ splice sites, and partially purified U1 would bind to both 3′ and 5′ splice site sequences on immobilized, single-stranded DNA or RNA, further purification of U1 by column chromatographic procedures showed binding only to immobilized 5′ splice sites.[278] Purified U1 bound to globin pre-mRNA transcripts, and RNAase T1 digestion of the bound transcript showed a protected region of several nucleotides comprising the 5′ splice site and part of the intron.[274] The majority of the 5′ splice site binding by U1 was abolished when the U1 was treated with protease K prior to the binding and immunoprecipitation assay, indicating that protein interactions and complementary base pairing are both important for the association of U1 with the 5′ splice site.[274] In 35S splicing intermediates or fully assembled spliceosomes, a larger region surrounding the 5′ splice site is protected from nuclease digestion, suggesting that the arrangement of the U1 snRNP shifts or that other factors associate with the substrate.[143]

Base pairing of the six nucleotides from position 2 to 8 of the 5′ terminus of U1 with a complementary oligonucleotide inhibited splicing substantially by itself, and site-directed cleavage of the double-stranded region with RNAase H completely abolished splicing *in vitro* extracts.[272] Degradation of the 5′ terminus of U1 results in loss of nuclease protection of the region encompassing the 5′ splice site and nearby exon sequences in *in vitro* splicing systems.[263] Not only does cleavage of the 5′ terminal nucleotides of U1 completely abolish splicing, but it reduces the association of U2, U4/U6, and U5 with the pre-mRNA as measured by immunoprecipitation with anti-Sm antibodies, indicating an important role for the U1 snRNP in the initial assembly of the spliceosome.[218,237,273,276] Surprisingly, the association of U1 with the 5′ splice site continues at a reduced level even after the entire 5′ sequences of U1 is removed by RNAase H cleavage, indicating that base pairing between the 5′ end of U1 and the 5′ splice junction is not the only interaction with the splicing substrate.[259] Unlike the other snRNPs, binding of U1 to pre-mRNA does not require ATP and, in fact, is inhibited by high concentrations of ATP.[67,259,260]

Further evidence for the role of base pairing between the 5′ splice site and the 5′ end of U1 was provided by Zhuang and Weiner,[277] who studied *in vivo* splicing in HeLa cells transfected with mutant adenovirus E1A transcripts and mutant U1. An E1A mutant with changes at the 5 and 6 positions of the 5′ splice site is defective for splicing, but if cotransfected with a mutant U1 that has complementary base pair changes at the 5′ end to compensate for the mutations at the 5′ splice junction of the transcript (and thus restore normal base pairing between U1 and the E1A transcript at the altered nucleotides), the altered E1A transcript is spliced normally. Thus, splice site mutations can be suppressed by mutations at the 5′ end of U1 that restore complementary base pairing. Not all 5′ splice site mutations can be suppressed by compensatory changes in U1, however, suggesting that base pairing between U1 and the 5′ splice site is necessary for splicing, but it is not the only factor involved.[277] A rigorous analysis of site-specific mutations in the region of the 5′ splice site was also consistent with the necessity of base-pairing between the U1 snRNA and the 5′ splice site for spliceosome function.[214,242]

The yeast yU1 functions in a similar manner to U1 in mammalian systems despite its size of 569 nucleotides. YU1 rapidly associates with the splicing substrate in an ATP-independent reaction within 30 s of mixing.[61,62,282] Like mammalian U1, base pairing between the 5′ end of yU1 and the 5′ splice junction is essential for its function. Cleavage of the 5′ end of yU1 using RNAase H and complementary oligonucleotides abolishes splicing.[58] Site-specific mutation of the splicing substrate and affinity purification of the splicing complexes indicate that in yeast, sequences surrounding both the 5′ splice site and the branch point are essential for the initial binding of the yU1 snRNP to the substrate.[45]

Mutations in the 5′ splice site of a yeast-splicing substrate substantially reduce, but do not totally abolish, U1 snRNP binding. Surprisingly, a single base pair substitution adjacent to the branch point completely abolished the binding of the yU1 snRNP to the substrate and subsequent assembly of the spliceosome.[45] This suggests that U1 has an important early step in spliceosome assembly that involves recognition of both the 5′ and 3′ splice sites. Observations in mammalian cell extracts that U1 snRNP will bind to substrates lacking a 5′ splice site also supports the notion that U1 snRNP binds to sites other than just the 5′ splice site.[219]

Several investigators have reported that monospecific U1 antiserum will immunoprecipitate a small amount of U2 and similarly monospecific U2 antiserum with immunoprecipitate some U1 snRNP.[165,239] These data suggest that there is interaction between U1 and U2 even in the absence of spliceosome assembly. Taken together, this suggests a vital role for the U1 snRNP in the assembly of the spliceosome. The data identify U1 snRNP recognition of sequences near the branch point and at the 5′ splice site, and suggest a crucial role for the U1 snRNP in the initial alignment of the different splice sites and in directing the subsequent assembly of the other snRNPs to form the mature spliceosome.

### a. U2

Psoralen cross-linking studies demonstrate that the U2 is hydrogen bonded to pre-mRNA in the cell nucleus.[237] Experimental analysis of the role of the U2 snRNP in the spliceosome suggests that it base pairs with conserved sequences near the

branch point in the intron, and that the fixed distance between the branch points and the 3' splice site and other specific factors that interact with the 3' splice site are important for U2 snRNP binding. Analysis of consensus branch point sequences in organisms ranging from yeast to humans showed that these sequences are complementary to a single-stranded region in the 5' half (33 to 38) of U2.[258,270] This is an attractive model, because the adenosine at the branch point would be bulged from the double helix, in a configuration that favors formation of the 5' to 2' linkage that occurs in the lariat splicing intermediate.[259] Several observations suggest that the U2 snRNP recognizes the branch point.

Studies in *in vitro* splicing systems indicate that an early step in splicing is the ATP-dependent association of a heat- and RNAase-sensitive factor with the branch point.[262] In splicing extracts, the branch point can be immunoprecipitated as a series of short fragments, protected from RNAase T1 digestion by use of U2-specific antibodies, and depletion of U2 snRNP from the extract eliminates the protection.[263] With a beta-globin substrate, in the 35S splicing complexes containing U2, both the branch point and the 3' splice site are protected from T1 RNAase digestion.[259] A similar protection is seen in complexes treated with heparin to remove the U1 snRNP, suggesting the critical role of U2 snRNP in protecting these sequences.[259] However, the exact regions of the splicing substrate protected by the U2 snRNP differ in detail depending on the sequence of the substrate.[263,270] With an adenovirus substrate, a region 6 to 20 nucleotides upstream of the branch point, but not the branch point itself, is protected from digestion with RNase A.[262] This may reflect different activities of the RNase A and RNase T1, or it may suggest that the exact interaction of the U2 snRNP with the substrate can vary, depending on the sequence.[259,280] Unlike the situation with U1, antibodies against U2 have no effect on splicing, which suggests the presence of these extra determinants on the U2 snRNP does not interfere with its function.[272]

Oligonucleotide-directed RNase H cleavage of U2 also provides strong evidence for the role of U2 in splicing. Cleavage of the 5' end of U2, and the loop region nearest the 3' end, abolish the formation of any splicing intermediates, while oligonucleotides that did not cause cleavage of U2 have no effect.[245,262,263,226] The 5' end of U2 is specifically required for nuclease protection of the branch point and nearby sequences.[245]

In a manner analogous to the experiments done with U1, a compensatory base change in U2 that restores base-pairing can suppress a mutation in the branch site that eliminates splicing.[270,6] This clearly indicates that nucleotides 33-38 of U2 can base-pair with the pre-mRNA substrate at the branch point but it may not be necessary because not all branch sites are complementary to U2.

Studies in *Xenopus* oocytes, where an injected snRNA gene is transcribed and assembled into functional snRNPs, suggest that the 3' end of the U2 snRNA that binds the U2-specific proteins is not essential for the function of the particle in pre-mRNA splicing.[266]

Mutational analysis suggests that the 3' splice site consensus sequences is more important for U2 snRNP binding than the branch point sequence itself. If the 3' splice site is mutated, U2 snRNP binding to the substrate and formation of the 35S splicing intermediate and subsequent splicing is substantially reduced or eliminated.[245,258,262,279] However, if the branch point sequences are mutated, a cryptic branch point will be activated that is the appropriate distance upstream of the 3' splice site.[262,279] This suggests a critical role for the 3' splice site or other factors associated with the 3' splice site in the binding of U2 snRNP to the splicing complex. Two factors have been identified that interact with the 3' splice site and are strong candidates for involvement in U2 binding to the branch point.

Highly purified U2 snRNPs will not bind to the branch point unless an additional protein component, the U2 auxiliary factor (U2AF), is added to the splicing reaction.[264] The U2AF binds specifically to the 3' splice site and is a necessary prerequisite to U2 snRNP binding to the branch point in a reconstituted system. Mutations of the polypyrimidine tract upstream of the 3' splice site or the 3' splice site itself block binding of U2AF to the splicing substrate and subsequent assembly of U2 snRNP. This suggests that the interaction of U2 snRNP with both the branch point sequences and the U2AF bound to the 3' splice site is critical for the stable association of U2 with the spliceosome.[264] A second protein, the U5-associated, 100-kDa protein, also binds to the 3' splice site and protects it from nuclease digestion.[132,134] The 100-kDa protein is clearly not the U2AF, and the role of the 100-kDa protein in U2 snRNP binding and spliceosome assembly is still not clear.

In yeast, yU2 plays essentially the same role in spliceosome formation and splicing as U2 in higher eukaryotes. Although yU2 at 1175 nucleotides is far larger than the 189-nucleotide U2, the sequence homology with U2 is primarily at the 5' 130 nucleotides, with a small region at the 3' end that is weakly homologous to U2.[65,66] Deletion of nearly all of the internal sequences (from 123 to 1082), which show regions of homology to U4, U5, and U6, results in normal splicing and no change in cell growth.[63,64] The internal sequences therefore do not have an essential function, although it is possible that they may be involved in interactions with other snRNPs and assist in assembly or stabilization of the spliceosome. Like U2, yU2 rapidly associates with the assembling spliceosome at an early stage, and stays bound throughout the splicing process.[253] YU2 is first detected with the formation of band III, the yeast analog of the 35S presplicing complex in mammalian systems, and deletion of yU2 sequences complementary to the branch point block splicing.[253,281] Unlike the observation in mammalian cells, the 3' splice site is not essential for the binding of yU2 to the splicing substrate and the assembly of the spliceosome.[281] This suggests that there is no equivalent of the mammalian U2AF in yeast.

### C. U4/U6 AND U5

U4, U5, and U6 are the last snRNPs to assemble with the spliceosome to form the mature 50S particle, and it is likely they

assemble with the substrate by protein-mediated interactions, rather than by complementary base-pairing with intron or splice junction sequences. Native gel electrophoresis suggests the U4/U6 and U5 snRNPs preassemble into a single 25S U4/U5/U6 particle in the nucleoplasm before binding to the 35S splicing complex containing U1 and U2 to form the mature 50-50S spliceosome complex.[180,181,237] U4 and U6 are stable components of spliceosomes isolated both by biotin-avidin affinity chromatography or by electrophoresis on native gels.[236,237] Kinetic studies demonstrate that U4/U5/U6 association with the spliceosome is a late step that occurs after the binding of U1 and U2 and is required for the splicing reaction.[180,237]

The 25S U4/U5/U6 particle is dynamic, with the majority of these three snRNPs in the 25S particle in nucleus, but the U5 snRNP dissociates under experimental conditions optimal for *in vitro* splicing.[244] The formation of the 25S particle is favored by millimolar ATP, but other forms of regulation may well be involved in 25S particle assembly.[368] U4 and U6 form a relatively stable particle with definite intermolecular base-pairing. In spliceosomes analyzed by native gel electrophoresis, U4 is lost from the splicing complex prior to 3' splice site cleavage. However, affinity purification procedures with complementary oligonucleotides find both U4 and U6 associated with the spliceosomes after completion of the splicing reaction.[289a] This suggests that the structure of the U4/U6 particle may change during the splicing reaction to destabilize U4 so that it is lost during native gel electrophoresis. In yeast, a 260-kDa protein has been identified that is part of the U5 snRNP and that is necessary for the ATP-dependent assembly of the U4/U5/U6 particle.[126]

Oligonucleotide-directed cleavage of U5 has not been successful; however, digestion of U4 or U6 at two different sites each inhibited splicing and formation of splicing intermediates in *in vitro* splicing extracts. Cleavage of the 5' end or the second loop of U4 (nucleotides 65 to 84) blocks formation of the 50S spliceosome from the 35S prespliceing complex.[365,366] Degradation of two regions of the second stem loop of U6, but not a single-stranded region, also inhibited spliceosome formation.[365,366] Other approaches suggest a critical role of the U5-specific, 100-kDa protein in the spliceosome.

The 100-kDa U5-associated protein specifically binds 3' splice site sequences immobilized on nitrocellulose, and protects 3' sequences from RNAase T1 digestion in splicing extracts.[132,134] Addition of oligonucleotides complementary to the polypyrimidine tract blocks binding of this protein to immobilized 3' splice sites, while competing complementary oligonucleotides to the 3' AG decrease, but do not abolish, binding of this protein.[132] The 100-kDa 3' binding protein could be immunoprecipitated with anti-SM, but not anti-(U1)RNP, antibodies, and comigrates with U5 in DEAE Sepharose chromatography.[132] The protein itself also binds anti-SM antibodies on immunoblots, indicating shared epitopes with the B and D snRNP proteins.[134] Its binding to U5 is relatively weak and/or transient, since it is abolished upon exposure to high $Mg^{++}$, resulting in its inability to be immunoprecipitated with anti-trimethylguanosine antibodies.[134] The role of this protein in the binding of the U4/U5/U6 complex or other snRNPs to the substrate is still not known.

Several observations suggest that the U4/U5/U6 snRNPs interact primarily with the U1 and U2 snRNPs and do not contact the pre-mRNA. In heparin-resistant splicing complexes that lack U1, there is no difference in the nuclease digestion patterns in spliceosomes containing or lacking U6 and U6 .[236] Also, in substrates lacking a 5' or 3' splice site that do not bind either U1 or U2 snRNP, respectively, the U4/U5/U6 snRNP binds inefficiently.[359] These observations, combined with the kinetic studies that show the U4/U5/U6 particle assembles into the spliceosome after the U1 and U2 particles, supports the notion that the U4/U5/U5 particle makes contact with the U1 and U2 snRNPs and possibly helps bring them in a stable particle. Although the site of U4/U5/U6 binding is not known for certain, nucleotides 65 to 77 of U4 could base-pair to nucleotides 102 to 108 of U1.

### *4. Splicing Summary*

Taken together, these data demonstrate the vital role of the five nucleoplasmic snRNPs in pre-mRNA splicing (Figure 5). The U1 snRNP is the first snRNP to bind to the substrate in an ATP-independent reaction that recognizes the 5' splice site and information in the branch point. This may have a critical role of aligning the splice sites for the subsequent transesterification reactions. Other factors then bind to the substrate at the 3' splice site that facilitate the binding of the U2 snRNP to the branch point. The U1 snRNA base-pairs with the conserved sequence motifs at the 5' splice site and the U2 snRNA base-pairs with sequences surrounding the branch point. A preassembled particle of U4/U5/U6 then assembles with the substrate through contacts with the U1 and U2 snRNP particles to form the mature spliceosome (Figure 5). The 5' splice site is cleaved and ligated to the branch point, and then the 3' site is cleaved and ligated to the 5' splice site. The spliceosomal components then disassemble to recycle for reuse.

Although the complete activities of the individual snRNPs in splicing reaction are not understood, observations suggest their functions include a structural role in the splicing complex. The occurrence of type II self-splicing introns in organelles, where introns are spliced by a similar sequence without the assistance of snRNPs, indicates that in some circumstances the functions provided by the snRNPs can be provided by the substrate itself.[229] However, it is likely this poses enormous constraints on the sequence of the substrate. When activities are supplied in *trans* by the snRNPs, this eliminates many of the constraints on the substrate and it is free to evolve in other directions. The recognition of specific sequence motifs in the substrate suggests the snRNPs provide a scaffolding that helps align the substrate for the transesterification reactions. The growing number of examples where RNA itself has enzymatic activity suggests that both the RNA and protein components of the snRNPs could be essential cofactors for the actual transesterification reactions.

Also, the use of snRNPs for splicing provides enormous opportunities for regulating the splicing reaction through modulation of the snRNPs. The observation that the snRNP-specific proteins have dynamic activities, independent of the core proteins, and the presence of a large number of low-abundance variant snRNAs with slightly different sequences are examples of how the activities of the individual snRNP particles can be regulated.

The snRNPs may also provide a mechanism for localizing pre-mRNA processing and transport to specific regions of the nucleus. The ultrastructural studies demonstrate the snRNPs are localized in specific regions of the nucleoplasm where they cluster into discrete structures. These may represent specific sites for both processing and export of pre-mRNA. Although many snRNPs appear to be soluble in the nucleoplasm, the enrichment of snRNPs in the nuclear matrix suggests the processing complexes occur in association with specific structures within the nucleus. These attachments could be directed by either interactions of the substrate pre-mRNAs or possibly by the snRNP particles. The tenfold lower abundance of the U4/U6 and U5 snRNPs compared with the U1 and U2 snRNPs, when they are all needed in single copies in the spliceosome, may reflect the different affinities of the snRNPs for the splicing complex. The reconstitution of *in vitro* splicing reactions from purified components will help identify the functions of the individual components in the splicing reaction. However, studies of the roles of snRNPs in localization and transport of pre-mRNA will require studies on intact nuclei.

## B. 3′ End Processing of Pre-mRNA

Pre-mRNAs in eukaryotic cells are transcribed as longer precursors. In addition to the removal of introns by RNA splicing as discussed above, the 3′ ends are also processed. Most mRNAs have a poly-A stretch of approximately 150 nucleotides added after 3′ cleavage, before the mature mRNA is exported to the cytoplasm. A small subset of the mRNAs, predominantly the histone mRNAs, are exported without the addition of poly-A nucleotides. Conserved sequence motifs that are essential for the specificity of the processing reactions are found surrounding the 3′ ends of both classes of pre-mRNAs. The histone transcripts have an essential stem-loop of approximately six base-pairs several nucleotides upstream from the 3′ end, and a highly conserved 8 nucleotide motif 13 to 15 nucleotides downstream of the 3′ end. Both are required for proper 3′ end cleavage. The polyadenylated mRNAs have a motif of AAUAAA approximately 10 to 20 nucleotides upstream from the polyadenylation site, and less well-conserved G- and U-rich sequence motifs that are within 50 nucleotides of the 3′ end.[31,283] SnRNPs are now identified as essential components of the processing complexes that generate the 3′ ends. In analogy to the functions of the snRNPs in formation of the spliceosome, base-pairing between the snRNA and conserved sequence motifs in the substrate are essential steps in the processing of histone transcripts and possibly the polyadenylation reaction as well.

### 1. Histone Pre-mRNA Processing

The U7 snRNP is a necessary cofactor for the 3′ cleavage of the histone pre-mRNA. U7 snRNP was the first snRNP identified on the basis of a functional assay. Cloned copies of a sea urchin histone H3 transcript injected into oocytes show little accurate 3′ end processing (most are extended at the 3′ end) unless sea urchin poly-A RNA was also injected.[285] Fractionation of the RNAs by sucrose gradient centrifugation and gel electrophoresis identify a 12S factor containing a 60-nucleotide RNA as the active termination factor. This factor, the U7 snRNP, is enriched by immunoprecipitation of sea urchin extracts with Sm antibodies and is approximately 3% the abundance of U1.[44] Sequencing of U7 indicates it is 57 nucleotides long, with a diagnostic trimethylguanosine cap, and sequences at the 5′ end of the snRNA can base-pair with the conserved sequence motif downstream from the histone 3′ end.[44] Although U7 does not contain a canonical sequence motif for binding the snRNP core proteins, mutational analysis suggests that a sequence of AG(Py)$_5$AAG between nucleotides 9 and 20 is responsible for binding the Sm reactive core proteins either directly or indirectly.[283]

The essential role of base-pairing between the U7 snRNA and the conserved downstream sequence element in the histone pre-mRNA was demonstrated in the oocyte system by compensatory base changes. Mutations in the conserved downstream sequence motif CAAGAAAGA in the histone H3 transcript that inhibited normal 3′ end formation could be rescued by compensatory changes in U7 that restore full base pairing of U7 to the sequence element.[283] Sequences in the terminal hairpin loop of mRNA can also base-pair with the U7 snRNA. However, mutational analysis suggests base-pairing with the stem loop is not vital to the function of the U7 snRNP.[284]

A mammalian *in vitro* histone pre-mRNA processing system, using cloned mammalian histone transcripts and a HeLa nuclear extract, has characterized several of the steps in the histone 3′ processing.[285] Accurate 3′ end processing requires both the U7 snRNP and a heat- and protease-sensitive factor that is not precipitable by anti-Sm antibodies,[45] in the *in vitro* system, anti-Sm antibodies precipitate RNAase T1-resistant fragments containing both the 3′ stem-loop of the histone mRNA and the 3 downstream sequence motif. Additional protected fragments are detected 30 min further into the processing reaction, including the 5′ cap of the histone mRNA and sequences further downstream from the 3′ end.[285] The factor binding the downstream sequence motif, but not to the stem-loop, is nuclease sensitive. This suggests that the factor binding the stem-loop is possibly a protein reactive with the Sm antiserum, but not a snRNP.[285] Cleavage of the 3′ end of U7 snRNA with complementary oligonucleotides and RNAase H abolishes histone pre-mRNA 3′ processing.[45,46,283] In mammalian cells the abundance of U7 is approximately 0.2% the abundance of U1.

Although absolutely conserved in sea urchin histone mRNA precursors, the 3′ downstream sequence element that base-pairs with U7 is less conserved in other organisms, with a consensus of

G/AAAAGA.[45] Sequence analysis of human U7 shows 50% sequence and considerable secondary structural homology with sea urchin U7, although the region of complementarity to the downstream processing element of histone pre-mRNA is ten nucleotides instead of the six nucleotides for sea urchin U7. The human U7 has several U residues in the putative base-pairing element that can base-pair with either A or G and allow greater flexibility in the recognition site.[45] The failure of *Xenopus* oocyte U7 snRNP to properly process the injected sea urchin H3 histone transcripts, but to process several other histone transcripts, suggests considerable heterogeneity of U7 among various organisms.[183] There are also data suggesting there may be several U7 snRNAs that could provide the possibility for alternative processing pathways.[183]

### 2. Polyadenylation

With limited exception, cytoplasmic mRNAs are polyadenylated in the cell nucleus before export to the cytoplasm. This reaction involves the endonucleolytic cleavage of the pre-mRNA at a specific position and the addition of approximately 200 adenylate residues at the 3' end. Data are accumulating that the U11 snRNP is part of the processing complex. However, base-pairing between a snRNA and the conserved sequence motifs has not been identified.

In *in vitro* polyadenylation systems, the AAUAAA sequence and in some circumstances fragments extending to the conserved downstream sequence elements are protected from RNAase cleavage and can be immunoprecipitated by Sm antisera.[4,7,288,289] When the complexes are cross-linked by UV irradiation, a nuclease-resistant fragment is immunoprecipitated by anti-Sm, but not anti-trimethylguanosine cap antisera.[290] This suggests that, like the stem-loop binding activity found in histone pre-mRNA processing, a Sm-reactive protein that is not a snRNP may bind this region. The involvement of a RNA in the polyadenylation reaction was suggested by the sensitivity of the *in vitro* reaction to micrococcal nuclease digestion.[76]

The U4 snRNA has sequences complementary to the AAUAAA polyadenylation reaction, but degradation of these sequences in the U4 does not block polyadenylation, suggesting an alternative snRNP is required for the reaction.[260,291] Recently, the *in vitro* polyadenylation extract has been fractionated into three essential components: a cleavage factor, a poly-A polymerase, and a ribonucleoprotein particle that copurifies with the U11 snRNA.[76] However, U11 does not have sequences complementary to the conserved AAUAAA polyadenylation motif, and it is unlikely it recognizes the substrate by base-pairing.[1,26] The possibility that other, yet unidentified snRNPs are essential for the processing reaction still remains.

## C. Preribosomal RNA Processing

U3 and U8 are unique among the characterized mammalian snRNPs in their localization in the nucleolus.[1,2] U8 is precipitated by anti-Sm antisera, but U3 is not. However, U8 lacks a consensus Sm binding site and could possibly be precipitated by association with other Sm-containing particles.[1] Recent studies indicate that U3 has a completely different protein composition than the other major snRNPs. An autoimmune serum directed against the 34-kDa nucleolar protein, fibrillarin, selectively immunoprecipitates the U3 snRNP and six polypeptides ranging in molecular weight from 12.5 to 74 kDa.[35] Immunocytochemistry indicates that fibrillarin and U3 are restricted to the fibrillar regions of the nucleolus, where rRNA transcription and processing occur before accumulation of more mature ribosomal subunits in the granular regions of the nucleolus.[34,133,134]

U3 was the first snRNA implicated in RNA processing of nuclear RNA, because approximately 30% of the total U3 remains hydrogen bonded to 45S and 32S rRNA precursors after rigorous extraction procedures.[31,34] U3 can be psoralen cross-linked to high molecular weight rRNA *in vivo*; however, the exact regions of base-pairing between U3 and the rRNA substrate are not clear.[35,172,231,293] Nucleotides 159 to 168 of U3 are complementary to sequences at the junction of the 5.8S RNA and the internal transcribed spacer of 32S RNA, which is the precursor of both 28S and 5.8S rRNAs. However, although these regions are single stranded, they are protected from chemical modification and RNAase A, H, or T1 cleavage, which suggests they are not available for base-pairing.[35,172] A single-stranded, relatively unprotected region of U3 103-112 also shows complementarity to regions of the transcribed spacer at the 3' end of 28S rRNA and is another candidate for regions of U3 that base-pairs with 28S pre-rRNA's.[35,132] However, psoralen cross-linking suggests U3 also base-pairs with regions at the 5' end of the 28S rRNA.[35,228] Therefore, the regions of the pre-rRNA recognized by the U3 snRNA and its possible function in recognizing specific sites of rRNA processing remain to be determined.

In yeast there are six snRNAs in the nucleolus, although several can be deleted without affecting yeast viability.[26] This opens the possibility that multiple snRNPs may assemble into a processing structure analogous to the spliceosome in the nucleoplasm. Further studies will be required to determine the exact functions of U3, U8, and other nucleolar snRNPs in ribosomal RNA processing.

## D. Outlook

The snRNP particles are emerging as cofactors for a wide range of RNA processing activities in the cell nucleus. They are another indication of the ability of RNA to manage its own affairs. Although the contributions of the snRNPs to the enzymatic activities of the processing reactions is not known, the snRNPs provide specificity by recognizing conserved sequence motifs in the substrates and provide a scaffolding upon which the processing events take place.

SnRNPs evolved with the eukaryotic cell. SnRNP assembly requires that the snRNAs appear transiently in the cytoplasm where they acquire their conserved core of snRNP proteins before returning permanently to the interphase nucleus. In organ-

elies, RNA splicing occurs that is similar to that in the eukaryotic nucleus. However, it occurs without the aid of snRNPs, suggesting the snRNPs appeared later and freed the substrate from the rigorous constraints of self-splicing. The common core of snRNP proteins may provide the specificity for assembling the multi-snRNP complexes of the spliceosome and other processing events through homotypic interactions between the core proteins. The RNA-free snRNP proteins in the cytoplasm display a number of homotypic interactions that may be a reflection of similar activities in the mature particles.

The abundance and stability of the snRNPs suggest that they are present as both active and recycling inactive forms in the nucleus, analogous to the functional cycle of the ribosomes in the cytoplasm. It is not known if the active snRNPs in the nucleoplasm are in a soluble compartment or attached to the fibrous skeleton of the nucleus. Quite possibly they appear in both contexts, and RNA processing may occur on a solid substrate that is one link in a chain of processing and transport events. RNA splicing and 3' end processing are sequential events for most pre-mRNAs, and the coupling of these processes to the maturation of the different types of pre-mRNA remains to be determined.

RNA processing is emerging as a major arena of gene expression. Genes can be expressed as families of closely related proteins by alternative RNA splicing and utilization of alternative 3' ends. The contributions of the snRNPs to these choices are not understood. One testable hypothesis is that variant snRNPs of low abundance that differ from the major canonical sequence contribute to the selection of these alternative sites. In addition, the activities of the snRNPs themselves may be regulated through the dynamic activities of the snRNP-specific proteins. Although the snRNP core proteins are as stable as the snRNA, the U1 and U2 snRNP-specific proteins turn over rapidly and the 70-kDa, U1-specific protein is present on some, but not all, U1 snRNPs. The snRNP proteins are also unusual because they are major autoantigens in SLE. By understanding how these proteins differ from other intracellular proteins, we may learn what features of these polypeptides predispose them to becoming autoantigens.

The growing number of low-abundance snRNPs identified suggests that snRNPs may be involved in a number of additional unidentified RNA-processing events. As the functions of the major snRNP particles emerge, many new questions emerge about the regulation of their activities and the details of their life cycle. A critical analysis of the functions of the snRNPs will await the ability to reconstruct functional snRNPs *in vitro* from purified components. With the availability of cloned genes for the snRNP proteins and specific variant snRNAs, *in vitro* systems will be able to evaluate the contributions of specific snRNAs and particular protein domains to RNA processing. The sophisticated genetic approaches available in yeast and other simple eukaryotes will also be valuable for defining and identifying the functions of other components of the processing complexes. The 1990s will be an exciting decade for studying the snRNP particles.

## REFERENCES


1. Busch, H., Reddy, R., Rothblum, L., and Choi, C. Y., SnRNAs, snRNPs and RNA processing, *Annu. Rev. Biochem.*, 51, 617, 1982.
2. Brunel, C., Sri-Widada, J., and Jeanteur, P., SnRNPs and scRNPs in eukaryotic cells, *Prog. Mol. Cell. Biol.*, 9, 1, 1985.
3. Dreyfuss, G., Philipsen, L., and Mattaj, I. W., Ribonucleoprotein particles in cellular processes, *J. Cell Biol.*, 106, 1419, 1988.
4. Birnstiel, M. L., Ed., Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, New York, 1988.
5. Guthrie, C. and Patterson, B., Spliceosomal snRNAs, *Annu. Rev. Genet.*, 22, 387, 1988.
6. Reddy, R., Completion of small RNA sequences, *Nucleic Acids Res.*, 13(Suppl.), r135, 1986.
7. Reddy, R., Henning, D., and Busch, H., Primary and secondary structure of U6 small nuclear RNA, *J. Biol. Chem.*, 260, 10930, 1985.
8. Montzka, K. A. and Steitz, J. A., Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc., *Proc. Natl. Acad. Sci. U.S.A.*, 85, 8885, 1988.
9. Kramer, A., Fractionation of HeLa cell nuclear extracts reveals novel small nuclear ribonucleoprotein particles, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 8408, 1987.
10. Riedel, N., Wise, J. A., Swerdlow, H., Mak, A., and Guthrie, C., Small nuclear RNAs from Saccharomyces cerevisiae: Unexpected diversity in abundance, size, and molecular complexity, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8097, 1986.
11. Riedel, N., Wolin, S., and Guthrie, C., A subset of yeast snRNAs contains functional binding sites for the highly conserved Sm antigen, *Science*, 235, 328, 1987.
12. Heilung-Larsen, P., Bialomowicz, I., and Frederiksen, S., Synthesis of low molecular weight RNA components in cells with a temperature-sensitive polymerase II, *Biochim. Biophys. Acta*, 609, 201, 1980.
13. Kunkel, G. R., Maser, R. L., Calvet, J. P., and Pederson, T., U6 small nuclear RNA is transcribed by RNA polymerase III, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8575, 1986.
14. Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D., The capped U6 small nuclear RNA is transcribed by RNA Polymerase III, *J. Biol. Chem.*, 262, 75, 1987.
15. Spector, D. L., Schrier, W. H., and Busch, H., Immunoelectron microscopic localization of snRNPs, *Biol. Cell*, 49, 1, 1983.
16. Kastner, B. and Lührmann, R., Electron microscopy of U1 small nuclear ribonucleoprotein particles: shape of the particle and position of the 5' RNA stem-loop, *EMBO J.*, 8, 377, 1989.
17. Feeney, R. J., Sauterer, R. A., Feeney, J. L., and Zieve, G. W., Cytoplasmic assembly and nuclear accumulation of mature snRNP particles, *J. Biol. Chem.*, 264, 5776, 1989.
17a. Feeney, R. A. and Zieve, G. W., Nuclear exchange of the U1 and U2 snRNP specific proteins, *J. Cell Biol.*, 110, in press.
18. Hashimoto, C. and Steitz, J. A., U4 and U6 RNA coexist in a single small nuclear ribonucleoprotein particle, *Nucleic Acids Res.*, 12, 3283, 1984.
19. Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H., and Lührmann, R., Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing, *EMBO J.*, 3, 1357, 1984.
20. Schmaus, C., MacAllister, G., Ohnsorge, Y., Hardin, J. A., and Lerner, M. R., A comparison of snRNP-associated Sm- autoantigens: human N, rat N and human B/B', *Nucleic Acids Res.*, 17, 6777, 1989.
21. Reichlin, M., Measurement of antibodies Sm and nRNP by ELISA: clinical and serological correlations, in *Mixed Connective Tissue Disease and Anti-Nuclear Antibodies*, Kasukawa, R. and Sharp, G. C., Eds., Elsevier Science, Amsterdam, 1987, 85.

22. Lerner, M. R. and Steitz, J. A., Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 5495, 1979.
23. Mimori, T., Hinterberger, M., Pettersson, T., and Steitz, J. A., Autoantibodies to the U2 small nuclear ribonucleoprotein in a patient with scleroderma-polymyositis overlap syndrome, *J. Biol. Chem.*, 259, 560, 1984.
24. Reuter, R., Lehner, C. F., Nigg, E. A., and Luhrmann, R., A monoclonal antibody specific for snRNPs U1 and U2, *FEBS Lett.*, 201, 25, 1986.
25. Parker, K. A. and Steitz, J. A., Structural analyses of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base-pairing with pre-rRNA, *Mol. Cell. Biol.*, 7, 2899, 1987.
26. Reimer, G., Pollard, K. M., Penning, C. A., Ochs, R. L., Lischwe, M. A., Busch, H., and Tan, E. M., Monoclonal autoantibody from a F1 mouse and some human scleroderma sera target a M, 34,000 nucleolar protein of the U3 snRNP particle, *Arthritis Rheum.*, 30, 793, 1987.
27. Zieve, G. W. and Feeney, R. A., Cytoplasmic assembly and nuclear transport of the snRNP particles, *Prog. Mol. Subcell. Biol.*, 11, 51, 1989.
28. Mattaj, I. W., Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding, *Cell*, 46, 90, 1986.
29. Mattaj, I. W. and De Robertis, E. M., Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins, *Cell*, 40, 111, 1985.
30. Patton, J. R., Patterson, R. J., and Pederson, T., Reconstitution of the U1 small nuclear ribonucleoprotein particle, *Mol. Cell. Biol.*, 7, 4030, 1987.
31. Sharp, P. A., Splicing of messenger RNA precursors, *Science*, 235, 766, 1987.
32. Mattaj, I. W. and Reed, R., The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing, *Neuron*, 335, 673, 1987.
33. Prestayko, A. W., Tonato, M., and Busch, H., Low molecular weight RNA associated with 28s nucleolar RNA, *J. Mol. Biol.*, 47, 505, 1970.
34. Zieve, G. W. and Penman, S., Small RNA species of the HeLa cell: metabolism and subcellular localization, *Cell*, 8, 19, 1976.
35. Mowry, K. L. and Steitz, J. A., snRNP mediators of 3' end processing: functional folds?, *Trends Biochem Sci.*, 13, 447, 1988.
36. Christofori, G. and Keller, W., 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor, and a snRNP, *Cell*, 54, 875, 1988.
37. Lienhard, J. P., Sri-Widada, J., Brunel, C., and Jeanteur, P., Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4, and U5 small nuclear RNAs, *J. Mol. Biol.*, 162, 573, 1982.
38. Mystinski, E. and Branlant, C., The small nuclear RNAs of Drosophila, *J. Mol. Biol.*, 180, 927, 1984.
39. Zieve, G. W., Benecke, B., and Penman, S., Synthesis of two classes of small RNA species in vivo and in vitro, *Biochemistry*, 16, 4520, 1977.
40. Wooley, J. C., Cone, R. D., Tartof, D., and Chung, S. Y., Small nuclear ribonucleoprotein complexes of Drosophila melanogaster, *Proc. Natl. Acad. Sci. U.S.A.*, 79, 6768, 1982.
41. Bruzik, J. P., Van Doren, K., Hirsh, D., and Steitz, J. A., Trans-splicing involves a novel form of small nuclear ribonucleoprotein particles, *Nature*, 335, 559, 1988.
42. Thomas, J. D., Conrad, R. C., and Blumenthal, T., The *C. elegans* trans-spliced leader RNA is bound to Sm and has a trimethylguanosine cap, *Cell*, 54, 533, 1988.
43. Van Doren, K. and Hirsch, D., Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in Caenorhabditis elegans, *Nature*, 335, 556, 1988.
44. Sirua, K., Galli, G., Busslinger, M., and Birnstiel, M., The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing, *EMBO J.*, 3, 2801, 1984.
45. Mowry, K. L. and Steitz, J. A., Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNAs, *Science*, 238, 1682, 1987.
46. Soldati, D. and Schumperli, D., Structural and functional characterization of mouse U7 small nuclear RNA active in 3' processing of histone pre-mRNA, *Mol. Cell. Biol.*, 8, 1518, 1988.
47. Hashimoto, C. and Steitz, J. A., A small nuclear ribonucleoprotein associates with the AAUAAA polyadenylation signal in vitro, *Cell*, 45, 581, 1986.
48. Lee, S. L., Murphy, S. C., Trimble, J. J., Desrosiers, R. C., and Steitz, J. A., Four novel U RNAs are encoded by a herpesvirus, *Cell*, 54, 599, 1988.
49. Tollervey, D. and Mattaj, I. W., Fungal small nuclear ribonucleoproteins share properties with plant and vertebrate U snRNPs, *EMBO J.*, 6, 469, 1987.
50. Kiss, T., Toth, M., and Solymosy, F., Plant small nuclear RNAs: nucleolar U3 snRNA is present in plants: partial characterization, *Eur. J. Biochem.*, 152, 259, 1985.
51. Kiss, T., Antal, M., and Solymosy, F., Plant small nuclear RNAs. II: the complete primary and secondary structure of broad bean U2 RNA: phylogenetic and functional implications, *Nucleic Acids Res.*, 15, 1337, 1987b.
52. Kiss, T., Antal, M., and Solymosy, F., Plant small nuclear RNAs II: U6 RNA and a 4.5S-like RNA present in plant nucleoli, *Nucleic Acids Res.*, 15, 5e3, 1987.
53. Kiss, T., Jakab, G., Antal, M., Palfi, K., Kis, M., and Solymosy, F., Plant small nuclear RNAs. V. U6 is present in broad bean plants in the form of sequence variants and is base-paired with U6 RNA, *Nucleic Acids Res. (England)*, 16, 5426, 1988.
53a. Singh, R. and Reddy, R., Gamma-monomethyl phosphate ester: the cap structure in spliceosomal U6 snRNA, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 8280, 1989.
54. Wise, J. A., Tollervey, D., Maloney, D., Swerdlow, H., Dunn, E. J., and Guthrie, C., Yeast contains small nuclear RNAs encoded by single copy genes, *Cell*, 35, 743, 1983.
55. Guthrie, C., Finding functions for small nuclear RNAs in yeast, *Trends Biochem Sci.*, October, 430, 1986.
56. Tollervey, D., A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA, *EMBO J.*, 6, 4169, 1987.
57. Hughes, J. M. X., Konings, D. A. M., and Cesareni, G., The yeast homologue of U3 snRNA, *EMBO J.*, 6, 2145, 1987.
58. Kretzner, L., Rymond, B. C., and Rosbash, M., *S. cerevisiae* U1 RNA is long and has limited primary sequence homology to metazoan U1 snRNA, *Cell*, 50, 593, 1987.
59. Siliciano, P., Jones, M. H., and Guthrie, C., *Saccharomyces cerevisiae* has a U1-like small nuclear RNA with unexpected properties, *Science*, 237, 1484, 1987.
60. Ares, M., Jr., U2 RNA from yeast is unexpectedly large and contains homology to vertebrate U4, U5 and U6 small nuclear RNAs, *Cell*, 47, 49, 1986.
61. Pikielny, C. W. and Rosbash, M., Specific small nuclear RNAs are associated with yeast spliceosomes, *Cell*, 45, 869, 1986.
62. Ruby, W. S. and Abelson, J., An early hexameric role of U1 small nuclear ribonucleoproteins in spliceosome assembly, *Nature*, 242, 1028, 1988.
63. Igel, A. H. and Ares, M., Jr., Internal sequences that distinguish yeast from mammalian U2 snRNA are unnecessary for pre-mRNA splicing, *Nature*, 334, 450, 1988.
64. Shuster, E. O. and Guthrie, C., Two conserved domains of yeast U2 snRNA are separated by 945 nonessential nucleotides, *Cell*, 55, 41, 1988.
65. Brow, D. A. and Guthrie, C., Spliceosomal RNA U6 is remarkably conserved from yeast to mammals, *Nature*, 334, 213, 1988.
66. Siliciano, P. G., Brow, D. A., Roiha, H., and Guthrie, C., An essential snRNA from *Saccharomyces cerevisiae* has properties predicted for U4, including interaction with a U6-like snRNA, *Cell*, 50, 585, 1987.
67. Tollervey, D., Wise, J. A., and Guthrie, C., A U4-like small nuclear

RNA is dispensible in yeast. *Cell*, 35, 253, 1983.
68. Patterson, B. and Guthrie, C., An essential yeast snRNA with a U5-like domain is required for splicing in vivo, *Cell*, 49, 613, 1987.
69. Bally, M., Hughes, J., and Cesareni, G., SnR30: a new, essential small nuclear RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, 16, 5291, 1988.
70. Branson, R. A. and Weiner, A. M., Human U1 RNA pseudogenes may be generated by both DNA and RNA mediated mechanisms, *Mol. Cell. Biol.*, 2, 815, 1982.
71. Parry, H. D., Scherly, D., and Mattaj, I. W., "Snurpogenesis": The transcription and assembly of U snRNP components, *Trends Biochem. Sci.*, 14, 12, 1989.
72. Skuratski, J. M., Lund, E., Murphy, J. T., Siebberg, T. H., Burgess, R. R., and Dahlberg, J. E., Multiple elements upstream of the coding region are required for accumulation of human U1 RNA in vivo, *J. Biol. Chem.*, 259, 8345, 1984.
73. Mougin, M., Ares, M., and Weiner, A. M., Human U2 snRNA genes contain an upstream enhancer, *EMBO J.*, 5, 987, 1986.
74. Mattaj, I. W., Lienhard, S., Jiricny, J., and De Robertis, E. M., An enhancer-like sequence within the *Xenopus* U2 gene promoter facilitates the formation of stable transcription complexes, *Nature*, 316, 163, 1985.
75. Ares, M., Jr., Mangin, M., and Weiner, A. M., Orientation-dependent transcriptional activator upstream of a human U2 snRNA gene, *Mol. Cell Biol.*, 5, 1560, 1985.
76. Dahlberg, J. E. and Lund, E., *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, Birnstiel, M., Ed., 1988, 38.
77. Carbon, P., Murgo, S., Ebel, J.-P., Krol, A., Tebb, G., and Mattaj, I. W., A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II, *Cell*, 51, 71, 1987.
78. Hernandez, N., Formation of the 3′ end of U1 snRNA is directed by a conserved sequence located downstream of the coding region, *EMBO J.*, 4, 1827, 1985.
79. Hernandez, N. and Weiner, A. M., Formation of the 3′ end of U1 snRNA requires compatible snRNA promoter elements, *Cell*, 47, 249, 1986.
80. Neuman de Vegvar, H. R., Lund, E., and Dahberg, J. E., 3′ End formation of U1 snRNA precursors is coupled to transcription from snRNA promoters, *Cell*, 47, 259, 1986.
81. Krol, A., Carbon, P., Ebel, J.-P., and Appel, B., *Xenopus* tropicalis U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes, *Nucleic Acids Res.*, 15, 2463, 1987.
82. Kunkel, G. R. and Pederson, T., Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used, *Genes Dev.*, 2, 196, 1988.
83. Das, G., Henning, D., Wright, D., and Reddy, R., Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III, *EMBO J.*, 7, 512, 1988.
84. Bark, C., Weller, P., Zabielski, J., Janson, L., and Pettersson, U., A distant enhancer is required for polymerase III transcription of a U6 gene, *Nature*, 328, 356, 1987.
85. Murphy, J. T., Burgess, R. R., Dahlberg, J. E., and Lund, E., Transcription of a gene for human U1 snRNA, *Cell*, 29, 265, 1982.
86. Lobe, S. M. and Muralhoff, W. F., Synthesis of U1 RNA in isolated mouse cell nuclei: initiation and 3′-end formation, *Mol. Cell. Biol.*, 3, 4290, 1987.
87. Morris, G. F., Price, D. H., and Marzluff, W. F., Synthesis of U1 RNA in a DNA-dependent system from sea urchin embryos, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 3674, 1986.
88. Lund, E. and Dahlberg, J. E., In vitro synthesis of vertebrate U1 snRNA, *EMBO J.*, 8, 387, 1989.
89. Naylor, S., Zabel, R., Manser, T., Gesteland, R., and Sukaguchi, A., Localization of human U1 small nuclear RNA genes to band 1p 36.3 of chromosome 1 by in situ hybridization, *Somatic Cell Mol. Genet.*, 10, 303, 1984.
90. Van Arsdell, S. W. and Weiner, A. M., Human genes for U2 small nuclear RNA are tandemly repeated, *Mol. Cell. Biol.*, 4, 492, 1984.
91. Lindgren, V., Ares, M., Jr., Weiner, A. M., and Francke, U., Human genes for U2 small nuclear RNA map to a major adenovirus 12 modification site on chromosome 17, *Nature*, 314, 115, 1985.
92. Lund, E., Dahlberg, J. E., and Forbes, D. J., The two small embryonic U1 small nuclear RNAs of *Xenopus laevis* are encoded by a major family of tandemly repeated genes, *Mol. Cell. Biol.*, 4, 2580, 1984.
93. Yu, J.-C., Nash, M. A., Santiago, C., and Marzluff, W. F., Structure and expression of a second sea urchin U1 RNA gene repeat, *Nucleic Acids Res.*, 14, 9977, 1986.
94. Lund, E. and Dahlberg, J. E., Two genes for human U1 small nuclear RNA, *J. Biol. Chem.*, 259, 2013, 1984.
95. Manser, T. and Gesteland, R. F., Human U1 loci: genes for human U1 RNA have dramatically similar genomic environments, *Cell*, 29, 257, 1982.
96. Lund, E., Heterogeneity of human U1 snRNAs, *Nucleic Acids Res.*, 16, 3813, 1988.
97. Patton, J. G. and Nielsen, E. D., U1 precursors variant 3′ flanking sequences are conserved in human cells, *J. Cell Biol.* 106, 175, 1987.
98. Murphy, J. T., Skuzeski, J. M., Lund, E., Steinberg, T. H., Burgess, R. R., and Dahlberg, J. E., Functional elements of the human U1 promoter. III. Identification of five separate regions required for efficient transcription of the human U1 snRNA, *J. Biol. Chem.*, 262, 1795, 1987.
99. Montesa, N. M., Lobe, S. M., and Marzluff, W. F., Expression of a mouse U1B gene in mouse L cells, *Gene*, 36, 311, 1985.
100. Ach, R. A. and Weiner, A. M., The highly conserved U small nuclear RNA 3′ end formation signal is quite tolerant to mutation, *Mol. Cell. Biol.*, 7, 2070, 1987.
101. Mattaj, I. W., Zeller, R., Carrasco, A. E., Jamrich, M., Lienhard, S., and De Robertis, E. M., U snRNA gene families in *Xenopus laevis*, in *Oxford Surveys of Eucaryotic Genes*, Vol. 2, Maclean, N., Ed., Oxford University Press, Oxford, 1985, 121.
102. Westin, G., Zabielski, J., Hammarstrom, K., Monstein, H.-J., and Bark, C., Clustered genes for human U2 RNA, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 3811, 1984.
103. Mattaj, I. W. and Zeller, R., *Xenopus laevis* U2 snRNA genes: tandemly repeated transcription units sharing 5′ and 3′ flanking homology with other RNA polymerase II transcribed genes, *EMBO J.*, 2, 1883, 1983.
104. Suh, D., Busch, H., and Reddy, R., Isolation and characterization of a human U3 small nucleolar RNA gene, *Biochem. Biophys. Res. Commun.*, 137, 1153, 1986.
105. Bark, C., Weller, P., Zabielski, J., and Pettersson, U., Genes for human U6 small nuclear RNA, *Gene*, 50, 333, 1986.
106. Suh, D., Busch, H., and Reddy, R., Human U3 small nucleolar RNA genes are localized to the nucleoplasm, *Biochem. Biophys. Res. Commun.*, 143, 638, 1987.
107. Hoffman, M. L., Korf, G. M., McNamara, K. J., and Stumph, W. E., Structural and functional analysis of chicken U4 small nuclear RNA genes, *Mol. Cell. Biol.*, 6, 3910, 1986.
108. Lund, E. and Dahlberg, J. E., Differential accumulation of U1 and U4 small nuclear RNAs during *Xenopus* development, *Genes Dev.*, 1, 39, 1987.
109. Kazmaier, M., Tebb, G., and Mattaj, I. W., Functional characterization of *X. laevis* U5 snRNA genes, *EMBO J.*, 6, 3071, 1987.
110. De Lorenzi, M., Rohrer, U., and Birnstiel, M. L., Analysis of a sea urchin gene cluster coding for the small nuclear U7 RNA, a rare RNA species implicated in the 3′ editing of histone precursor mRNAs, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 3243, 1986.
111. Mattaj, I. W., Dathan, N. A., Parry, H. D., Carbon, P., and Krol, A., Changing the RNA polymerase specificity of U snRNA gene promoters,

*Cell*, 55, 435, 1988.
112. Hsu, G., Henning, D., and Reddy, R., Structure, organization, and transcription of Drosophila U6 small nuclear RNA genes, *J. Biol. Chem.*, 262, 1187, 1987.
113. Tani, T. and Ohshima, Y., The gene for the U6 small nuclear RNA in fission yeast has an intron, *Nature*, 337, 87, 1989.
114. Brietbart, R. E., Andreadis, A., and Nadal-Ginard, B., Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes, *Annu. Rev. Biochem.*, 56, 467, 1987.
115. Kato, N. and Harada, F., New U1 RNA species found in Trend SFFV (Spleen Focus Forming Virus)-transformed cells, *J. Biol. Chem.*, 260, 7215, 1985.
116. Lund, E., Kahan, B., and Dahlberg, J. E., Differential control of U1 small nuclear RNA expression during mouse development, *Science*, 239, 1271, 1985.
117. Lobo, S. M., Maratsos, W. F., Seufert, A. C., Dean, W. L., Schultz, G. A., Sumerly, C., and Schatten, G., Localization and expression of U1 RNA in early mouse embryo development, *Dev. Biol.*, 127, 340, 1988.
118. Forbes, D. L., Kirschner, M. W., Caput, D., Dahlberg, J. E., and Lund, E., Differential expression of multiple U1 small nuclear RNAs in oocytes and embryos of *Xenopus laevis*, *Cell*, 38, 681, 1984.
119. Forbes, D. J., Kornberg, T. B., and Kirschner, M. W., Small nuclear RNA transcription and ribonucleoprotein assembly in early Xenopus development, *J. Cell Biol.*, 97, 62, 1983.
120. Lund, E., Bostock, C. J., and Dahlberg, J. E., The transcription of *Xenopus laevis* embryonic U1 snRNA genes changes when oocytes mature into eggs, *Genes Dev.*, 1, 47, 1987.
121. Lerner, E. A., Lerner, M. R., Janeway, C. A., Jr., and Steitz, J. A., Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 2737, 1981.
122. Williams, D. G., Sheeks, M. R., Smith, P. R., and Muni, R. N., Murine lupus ribonucleoprotein autoantibodies define five epitopes on two different Sm polypeptides, *Immunology*, 58, 495, 1986
123. Billings, P. B., Barton, J. R., and Hoch, S. O., A murine monoclonal autobody recognizes the 13,000 molecular weight polypeptide of the Sm small nuclear ribonucleoprotein complex, *J. Immunol.*, 135, 428, 1985
124. Billings, P. B., Allen, R. W., Jensen, F. C., and Hoch, S. O., Anti-RNP monoclonal antibodies derived from a mouse strain with lupus-like autoimmunity, *J. Immunol.*, 128, 176, 1982.
125. Winkler, A., Watson-McKown, R., and Wise, K. S., Monoclonal autoantibody recognizing a unique set of small nuclear ribonucleoprotein complexes, *J. Immunol.*, 140, 69, 1988.
126. White, P. J. and Hoch, S. O., Definition of the antigenic polypeptides in the Sm and RNP ribonucleoprotein complexes, *Biochem. Biophys. Res. Commun.*, 102, 365, 1981.
127. White, P. J., Billings, P. R., and Hoch, S. O., Assays for the Sm and RNP autoantigens: the requirement for RNA and influence of the tissue source, *J. Immunol.*, 128, 2251, 1982.
128. Peterson, J., Hinterberger, M., Mimori, T., Gottlieb, E., and Steitz, J. A., The structure of mammalian small nuclear ribonucleoproteins: identification of multiple protein components reactive with anti-(U1)RNP and anti-Sm autoantibodies, *J. Biol. Chem.*, 259, 5903, 1984.
129. McAllister, G., Roby Shemkovitz, A., Amara, S. G., and Lerner, M. R., cDNA sequences of the rat U snRNP-associative protein N: description of a potential Sm epitope, *EMBO J.*, 8, 1177, 1989.
129a. Rokeach, L. A., Jannatipour, M., and Hoch, S., Heterologous expression of epitope mapping of a human Sm-B/B autoantigen, *J. Immunol.*, 145(3), 1990.
130. Craft, J., Mimori, T., Olsen, T. L., and Hardin, J. A., The U2 small nuclear ribonucleoprotein particle as an autoantigen. Analysis with sera from patients with overlap syndromes, *J. Clin. Invest.*, 81, 1716, 1988.
131. Habets, W. J., Hoet, M. H., Bringmann, P., Luhrmann, R., and van Venrooij, W. J., Autoantibodies to ribonucleoprotein particles containing U2 small nuclear RNA, *EMBO J.*, 4, 1545, 1985.
132. Sillekens, P. T., Habets, W. J., Beijer, R. P., and van Venrooij, W. J., cDNA cloning of the human U1 snRNA-associated A protein: extensive homology between U1 and U2 snRNP-specific proteins, *EMBO J.*, 6, 3841, 1987.
133. Lischwe, M. A., Ochs, R. L., Reddy, R., Cook, R. G., Yeoman, L. C., Tan, E. M., Reichlin, M., and Busch, H., Purification and partial characterization of a nucleolar scleroderma antigen (M<sub>r</sub>34 000; pI, 8, 5) rich in NG,NG-dimethylarginine, *J. Biol. Chem.*, 260, 14304, 1985
134. Ochs, R. L., Lischwe, M. A., Spohn, W. H., and Busch, H., Fibrillarin: a new protein of the nucleolus identified by autoimmune sera, *Biol. Cell.*, 54, 124, 1985.
135. Reuter, R. and Luhrmann, R., Immunization of mice with purified U1 small nuclear ribonucleoproteins (RNP) induces a pattern of antibody specificities characteristic of the anti-Sm and anti-RNP autoimmune response of patients with lupus erythematosus, as measured by monoclonal antibodies, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8549, 1986.
136. Oldstone, M. B., Molecular mimicry and autoimmune disease, *Cell*, 50, 819, 1987
137. Query, C. C. and Keene, J. D., A human autoimmune protein associated with U1 RNA contains a region of homology that is cross-reactive with retroviral recroviral p30gag antigen, *Cell*, 51, 211, 1987.
138. Rocheleau, M., Greenland, H., Valle, H., and Larsen, C. J., Human autoimmune serum antibodies against gag gene p30 retroviral protein also react with a U1-SnRNP 68K conjugate protein, *Biol. Cell*, 60, 71, 1987.
139. Mattaj, I. W., Habets, W. J., and van Venrooij, W. J., Monospecific antibodies reveal details of U2 snRNP structure and interaction between U1 and U2 snRNPs, *EMBO J.*, 5, 997, 1986
140. Lehrmann, R., Appel, B., Bringmann, P., Rinke, J., Reuter, R., Reuter, S., and Bald, R., Isolation and characterization of rabbit anti-2, 2, 7-trimethylguanosine antibodies, *Nucleic Acids Res.*, 10, 7103, 1982
141. Smith, J. H. and Ehrlich, G. L., Autoantibodies elicited by N2, N2,7-trimethylguanosine react with small nuclear RNAs and ribonucleoproteins, *J. Biol. Chem.*, 258, 6636, 1983
142. Bochnig, P., Reuter, R., Bringmann, P., and Luhrmann, R., A monoclonal antibody against the 2, 2, 7-trimethylguanosine cap that reacts with intact U small nuclear ribonucleoproteins as well as with 7-methylguanosine-capped RNAs, *Eur. J. Biochem.*, 168, 461 1987
143. Bringmann, P., Rinke, J., Appel, B., Reuter, R., and Luhrmann, R., Purification of snRNPs U1, U2, U4, U5 and U6, with 2, 2, 7-trimethylguanosine-specific antibody and definition of their constituent proteins reacting with anti-Sm and anti-(U1)RNP antisera, *EMBO J.*, 2, 1129, 1983b.
144. Bringmann, P. and Luhrmann, R., Antibodies specific for N6-methyladenosine react with intact snRNPs U2 and U4/U6, *FEBS Lett.*, 213, 309, 1987.
145. Crabbe, B. and Steitz, J. A., Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes, *Mol. Cell. Biol.*, 7, 281, 1987
146. Hinterberger, M., Peterson, L., and Steitz, J. A., Isolation of small nuclear ribonucleoproteins containing U1, U2, U4, U5, and U6 RNAs, *J. Biol. Chem.*, 258, 2604, 1983.
147. Kinlaw, C. S., Robberson, B. L., and Berget, S. M., Fractionation and characterization of human small nuclear ribonucleoproteins containing U1 and U2 RNAs, *J. Biol. Chem.*, 258, 7181, 1983.
148. Bringmann, P. and Luhrmann, R., Purification of the individual snRNPs U1, U2, U5, and U4/U6 from HeLa cells and characterization of their protein constituents, *EMBO J.*, 5, 3509, 1986
148a. Andersen, J., Feeney, R. J., and Zieve, G. W., Identification of the D' snRNP core protein, in preparation
149 Reuter, R., Rothe, S., and Luhrmann, R., Molecular relationships between U snRNP proteins as investigated by rabbit antisera and peptide mapping, *Nucleic Acids Res.*, 15, 4021, 1987

150. Fisher, D. E., Conner, G. E., Reeves, N. H., Wisniewolski, R., and Blobel, G., Small nuclear ribonucleoprotein particle assembly in vivo: demonstration of a 65 RNA free core precursor and posttranslational modification. *Cell*, 42, 751, 1985.

151. Sauterer, R. A., Goyal, A., and Zieve, G. W., Cytoplasmic assembly of snRNP particles from 6S and 20S RNA-free intermediates, *J. Biol. Chem.*, 263, 1048, 1990.

152. Task, J., Alibert, C., Temsamani, J., Revellaud, L., Cathala, G., Brunel, C., and Jeanteur, P., A protein that specifically recognizes the 3' splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein, *Cell*, 42, 755, 1985.

153. Lelay-Taha, M.-N., Revellaud, L., Sri-Widada, J., Brunel, C., and Jeanteur, P., RNA-protein organization of U1, U5, and U4-U6 small nuclear ribonucleoproteins in HeLa cells, *J. Mol. Biol.*, 189, 519, 1986.

154. Gerke, V., and Steitz, J. A., A protein associated with small nuclear ribonucleoprotein particles recognizes the 3' splice site of premessenger RNA, *Cell*, 47, 973, 1986.

155. Revellaud, L., Lelay-Taha, M. N., Sri-Widada, J., Brunel, C., and Jeanteur, P., Mg$^{2+}$ induces a sharp and reversible transition in U1 and U2 small nuclear ribonucleoprotein configurations, *Mol. Cell. Biol.*, 4, 1800, 1984.

156. Lossky, M., Anderson, G. J., Jackson, S. J., and Beggs, J., Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions, *Cell*, 42, 1015, 1985.

156a. Bjorn, S. P., Soltyk, A., Beggs, J. D., and Friesen, J. D., PRP4 (RNA4) from *Saccharomyces cerevisiae*: its gene product is associated with the U4/U6 small nuclear ribonucleoprotein particle, *Mol. Cell. Biol.*, 9, 3698, 1989.

157. Ohosone, Y., Mimori, T., Griffith, A., Akizuki, M., Homma, M., Craft, J., and Hardin, J. A., Molecular cloning of a Sm autoantigen: derivation of a cDNA for a B polypeptide of the U series of small nuclear ribonucleoprotein particles, *Proc. Natl. Acad. Sci. U.S.A.*, 264, 3026, 1989.

158. Rokeach, L. A., Junnatspour, E.-A. H., Hately, J. A., and Hoch, S. O., Primary structure of the human snRNP polypeptide as deduced by cDNA analysis, *J. Biol. Chem.*, in press.

159. McAllister, G., Amara, S. G., and Lerner, M. R., Tissue specific expression and cDNA cloning of small nuclear ribonucleoprotein-associated polypeptide N, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 5296, 1988

160. Rokeach, L. A., Hately, J. A., and Hoch, S. O., Molecular cloning of a cDNA encoding the human Sm-D autoantigen, *Proc. Natl. Acad. Sci U.S.A.*, 85, 4832, 1988.

161. Wieben, E. D., Rohleder, A. M., Nenninger, J. M., and Pederson, T., cDNA cloning of a human autoimmune nuclear ribonucleoprotein antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 7914, 1985.

162. Stanford, D. R., Lapeyre, B., Perry, C. A., Holicky, E. L., Harvey, S. E., Rothrister, N. M., Rohder, K., Jr., Luhrmann, R., and Wieben, E. D., The complete primary structure of the human snRNP E protein, *Nucleic Acids Res.*, 16, 10593, 1988.

163. Stanford, D. R., Perry, C. A., Holicky, E. L., Rohleder, A. M., and Wieben, E. D., The small nuclear ribonucleoprotein E protein gene contains four introns and has upstream similarities to genes for ribosomal proteins, *J. Biol. Chem.*, 263, 17727, 1988.

164. Theissen, H., Etzerodt, M., Reuter, R., Schneider, C. F., Lottspeich, F., Argos, P., Luhrmann, R., and Philipson, L., Cloning of the human cDNA for the U1 RNA-associated 70K protein, *EMBO J.*, 5, 3209, 1986.

164a. Query, C. C., Bentley, R. C., and Keene, J. C., A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein, *Cell*, 57, 89, 1989.

165. Yamamoto, K., Miura, H., Morai, Y., Yoshinoya, S., Goto, M., Nishioka, K., and Miyamoto, R., Isolation and characterization of a complementary DNA expressing human U1 small nuclear ribonucleoprotein C polypeptide, *J. Immunol.*, 140, 311, 1988.

166. Silbekens, P. T. G., Beijer, R. P., Habets, W. J., and van Venrooij, W. J., Human U1 snRNP-specific C protein: complete cDNA and protein sequence and identification of a multigene family, *Nucleic Acids Res.*, 16, 8307, 1988.

167. Habets, W. J., Silbekens, P. T. G., Hoet, M. H., Schalken, J. A., Roebroek, A. J. M., Leunissen, J. A. M., van de Ven, W. J. M., and van Venrooij, W. J., Analysis of a cDNA clone expressing a human autoimmune antigen: full-length sequence of the U2 small nuclear RNA-associated B" antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 2421, 1987.

168. Fresco, L. D., Buratlu, M. D., and Keene, J. D., Rapid inhibition of processing and assembly of small nuclear ribonucleoproteins after infection with vesicular stomatitis virus, *Mol. Cell. Biol.*, 7, 1148, 1987.

169. Fresco, L. D. and Keene, J. D., Biochemical exchange and molecular analysis of A' implicate a leucine periodicity in protein:protein interactions of the U2 snRNP, *Mol. Cell. Biol.*, in press.

170. Swanson, M. S., Nakagawa, T. Y., LeVan, K., and Dreyfuss, G., Primary structure of human nuclear ribonucleoprotein particle C protein: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA and pre-RNA-binding proteins, *Mol. Cell. Biol.*, 7, 1731, 1987.

171. Patton, J. R. and Pederson, T., The M, 70,000 protein of the U1 small nuclear ribonucleoprotein particle binds to the 5' stem-loop of U1 RNA and interacts with the Sm domain proteins, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 747, 1988.

172. Jeppesen, C., Stebbins-Boaz, B., and Gerbi, S. A., Nucleotide sequence determination and secondary structure of Xenopus U3 snRNA, *Nucleic Acids Res.*, 16, 2127, 1988.

173. Rinke, J., Appel, B., Digweed, M., and Luhrmann, R., Localization of a base-paired interaction between small nuclear RNAs U5 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen crosslinking, *J. Mol. Biol.*, 185, 721, 1985.

174. Harris, S. G., Matin, T. E., and Smith, H. C., Reversible chemical cross-linking and ribonuclease digestion: analysis of the organization of proteins in ribonucleoprotein particles, *Mol. Cell. Biochem.*, 84, 17, 1988.

175. Harris, S. G., Hoch, S. O., and Smith, H. C., Chemical cross-linking of Sm and RNP antigenic proteins, *Biochemistry*, 27, 4595, 1988.

176. Luhrmann, R., snRNP proteins in *Structure and Function of Major and Minor Small Nuclear Ribonucleoproteins*, Springer-Verlag, New York, 1988.

177. Liu, W.-J., and Pederson, T., Ribonucleoprotein organization of eukaryotic RNA, XXXI. Structure of the U1 small nuclear ribonucleoprotein, *J. Mol. Biol.*, 180, 947, 1984.

178. Hamm, J., Kazmaier, M., and Mattaj, I. W., In vitro assembly of U1 snRNPs, *EMBO J.*, 6, 3479, 1987.

179. Bringmann, P., Reuter, R., Rinke, J., Appel, B., Bald, R., and Luhrmann, R., 3'-terminal caps of snRNA are accessible for reaction with 2, 2, 7-trimethylguanosine-specific antibody in intact snRNPs, *J. Biol. Chem.*, 258, 2745, 1983.

180. Konarska, M. M. and Sharp, P. A., Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes, *Cell*, 49, 763, 1987.

181. Lamond, A. I., Konarska, M. M., Grabowski, P. J., and Sharp, P. A., Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein, *Proc. Natl Acad Sci U.S.A.*, 85, 411, 1988.

182. Eliceiri, G. L., Formation of low molecular weight RNA species in HeLa cells, *J. Cell. Physiol.*, 102, 195, 1980.

183. Chandrasekharappa, S. C., Smith, J. H., and Eliceiri, G. L., Biosynthesis of small nuclear RNAs in human cells, *J. Cell. Phys.*, 117, 159, 1983.

184. Eliceiri, G. L., Maturation of low-molecular-weight RNA species, in *The Cell Nucleus*, Vol. 8, Busch, H., Ed., Academic Press, New York, 1981, 307.

185. Rinke, J. and Steitz, J. A., Association of the lupus antigen La with a subset of U6 snRNA molecules, *Nucleic Acids Res.*, 13, 2617, 1985.

186. Madore, S. J., Wieben, E. D., and Pederson, T., Eukaryotic small

ribonucleoproteins: anti-La human autoantibodies react with U1 RNA-protein complexes. *J. Biol. Chem.*, 259, 1929, 1984.
187. Madore, S. J., Wieben, E. D., and Pederson, T., Intracellular site of U1 small nuclear RNA: processing and ribonucleotide assembly, *J. Cell Biol.*, 98, 188, 1984.
188. Zieve, G. W., Sauterer, R. A., and Feeney, R. J., Newly synthesized snRNA appear transiently in the cytoplasm, *J. Mol. Biol.*, 199, 250, 1988.
189. Gurney, T., Jr. and Eliceiri, G. L., Intracellular distribution of low molecular weight RNA species in HeLa cells, *J. Cell Biol.*, 87, 398, 1980.
190. Sauterer, R. A., Feeney, R. J., and Zieve, G. W., Cytoplasmic assembly of snRNP particles from stored proteins and newly transcribed snRNAs in L929 mouse fibroblasts, *Exp. Cell Res.*, 176, 344, 1988.
191. Madore, S. J., Wieben, E. D., Kunkel, G. R., and Pederson, T., Precursors of U4 small nuclear RNA, *J. Cell Biol.*, 99, 1140, 1984.
192. Zieve, G. W., Cytoplasmic maturation of the snRNAs, *J. Cell Physiol.*, 131, 287, 1987.
193. Kleinschmidt, A. M. and Pederson, T., Accurate and efficient 3′ processing of U2 small nuclear RNA precursor in a fractionated cytoplasmic extract, *Mol. Cell. Biol.*, 7, 3131, 1987.
194. De Robertis, E. M., Lienhard, S., and Parisot, R. F., Intracellular transport of microinjected 5S and small nuclear RNA, *Nature*, 295, 572, 1982.
195. Fisher, D. E., Conner, G. E., Reeves, W. H., Blobel, G., and Kunkel, H. G., Systemic and sensitivity of human small nuclear ribonucleoproteins generated by cell-free translation, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 6356, 1983.
196. Zeller, R., Neffenegger, T., and De Robertis, E. M., Nucleocytoplasmic distribution of snRNPs and stockpiled snRNA-binding proteins during oogenesis and early development in *Xenopus laevis*, *Cell*, 32, 425, 1983.
197. Fritz, A., Parisot, R. F., Newmeyer, D., and De Robertis, E. M., Small nuclear U-RNPs in *Xenopus laevis* development: uncoupled accumulation of the protein and RNA components, *J. Mol. Biol.*, 178, 273, 1984.
198. De Robertis, E. M., Nucleocytoplasmic segregation of proteins and RNAs, *Cell*, 32, 1021, 1983.
199. Konings, D. A. M. and Mattaj, I. W., Mutant U2 snRNAs of Xenopus which can form an altered higher order RNA structure are unable to enter the nucleus, *Exp. Cell Res.*, 167, 1, 1987.
200. Paine, P. L. and Horowitz, S. B., The movement of material between nucleus and cytoplasm, *Cell Biol.*, 4, 299, 1980.
201. Newmeyer, D. L. and Forbes, D. J., Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation, *Cell*, 52, 641, 1988.
202. Richardson, W. D., Mills, A. D., Dilworth, S. M., Laskey, R. A., and Dingwall, C., Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores, *Cell*, 52, 655, 1988.
203. Dingwall, C., Sharnick, S. V., and Laskey, R. A., A polypeptide domain that specifies migration of nucleoplasmin into the nucleus, *Cell*, 30, 449, 1982.
204. Kalderon, D., Richardson, W. D., Markham, A. F., and Smith, A. E., Sequence requirements for the nuclear localization of simian virus 40 large-T antigen, *Nature*, 311, 33, 1984.
205. Roberts, B. L., Richardson, W. D., and Smith, A. E., The effect of protein context on nuclear location signal function, *Cell*, 50, 465, 1987.
206. Davey, J., Dimmock, N. J., and Colman, A., Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes, *Cell*, 40, 667, 1985.
207. Deng, J. S., Takasaki, Y., and Tan, E. M., Nonhistone nuclear antigens reactive with autoantibodies. Immunofluorescent studies of distribution in synchronized cells, *J. Cell Biol.*, 91, 654, 1981.
208. Spector, D. L., Colocalization of U1 and U2 small nuclear RNPs by immunocytochemistry, *Biol. Cell.*, 51, 109, 1984.
209. Smith, H. C., Spector, D. L., Woodcock, C. L. F., Oehs, R. L., and Bhorjee, J., Alterations in ribosomin conformation are accompanied by reorganization of nonchromatin domains that contain U-snRNP protein p28 and nuclear protein p107, *J. Cell Biol.*, 101, 560, 1985.
210. Nyman, U., Hallman, H., Hadlaczky, G., Pettersson, I., Sharp, G., and Ringertz, N. R., Intranuclear localization of snRNP antigens, *J. Cell Biol.*, 102, 137, 1986.
211. Symington, J., Gurney, T., Jr., and Eliceiri, G. L., Subcellular location of polypeptides that react with anti-Sm and anti-RNP antibodies, *Biochem. Biophys. Res. Commun.*, 120, 81, 1984.
212. Spector, D. L., Watt, R. A., and Sullivan, N. F., The v- and c-myc oncogene proteins colocalize in cells with small nuclear ribonucleoprotein particles, *Oncogene*, 1, 5, 1987.
213. Fakan, S., Leser, G., and Martin, T., Immunoelectron microscope visualization of nuclear ribonucleoprotein antigens within spread transcription complexes, *J. Cell Biol.*, 103, 1153, 1986.
214. Pavian, E., Presence of RNA in interchromatin granules of cells treated with DRB, *Biol. Cell.*, 43, 147, 1982.
215. Beyer, A. L. and Osheim, Y. N., Splice site selection, rate of splicing and alternative splicing on nascent transcripts, *Genes Dev.*, 2, 754, 1988.
216. Sass, H. and Pederson, T., Transcription-dependent localization of U1 and U2 small nuclear ribonucleoproteins at major sites of gene activity in polytene chromosomes, *J. Mol. Biol.*, 180, 911, 1984.
217. Harris, S. G. and Smith, H. C., snRNP core protein enrichment in the nuclear matrix, *Biochem. Biophys. Res. Commun.*, 152, 1382, 1988.
218. Smith, H. C., Oehs, R. L., Fernandez, E. A., and Spector, D. L., Macromolecular complex containing nuclear protein p107 and U-snRNP protein p28: further evidence for an in situ nuclear matrix, *Mol. Cell. Biochem.*, 70, 151, 1986.
219. Verheijen, R., Kuijpers, H., Vooijs, P., van Venrooij, W., and Ramaekers, F., Distribution of the 70K U1 RNA-associated protein during interphase and mitosis, *J. Cell Sci.*, 86, 173, 1986.
220. Zieve, G. W. and Penman, S., Subnuclear particles containing a small nuclear RNA and hnRNA, *J. Mol. Biol.*, 145, 501, 1981.
221. Pederson, T. and Bhorjee, J. S., Evidence for a role of RNA in eukaryotic chromosome structure. Metabolically stable, small nuclear RNA species are covalently linked to chromosomal DNA in HeLa cells, *J. Mol. Biol.*, 128, 451, 1979.
222. Zieve, G. W. and Sliwka, S. B., Removal of cellular water prevents the reformation of the interphase nucleus, *J. Cell. Physiol.*, 129, 85, 1986.
223. Reuter, R., Appel, B., Rinke, J., and Luhrmann, R., Localization and structure of snRNPs during mitosis, *Exp. Cell Res.*, 159, 63, 1985.
224. Spector, D. L. and Smith, H. C., Redistribution of U-snRNPs during mitosis, *Exp. Cell Res.*, 160, 87, 1986.
225. Nash, M. A., Kozak, S. E., Angerer, L. M., Angerer, R. C., Schaffer, H., Schatten, G., and Marshall, W. F., Sea urchin maternal and embryonic U1 RNAs are spatially segregated in early embryos, *J. Cell Biol.*, 104, 1133, 1987.
226. Ruzsdjic, S. and Pederson, T., Evidence for an association between U1 RNA and interspersed repeat single-copy RNAs in the cytoplasm of sea urchin eggs, *Develop.*, 101, 109, 1987.
227. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. R., and Sharp, P. A., Splicing of messenger RNA precursors, *Annu. Rev. Biochem.*, 55, 1119, 1986.
228. Steitz, J. A., Black, D. L., Gerke, V., and Parker, K. A., Functions of the abundant U-snRNPs, in *Small Nuclear Ribonucleoprotein Particles: Structure and Functions*, Birnstiel, M. L., Ed., Springer-Verlag, Heidelberg, 1987, 115.
229. Cech, T. R., The generality of self-splicing RNA: relationship to nuclear mRNA splicing, *Cell*, 44, 207, 1986.
230. Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R., Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vivo, *Cell*, 38, 317, 1984.
231. Padgett, R. A., Konarska, M. M., Grabowski P. J., Hardy, S. F., and Sharp, P. A., Lariat RNAs as intermediates and products in the splicing

of messenger RNA precursors, *Science* 225, 898, 1984.
232. DiMario, P., Kaltwasser, G., and Goldenberg, G., Partial purification and properties of a pre-mRNA splicing activity, *J. Biol. Chem.*, 260, 1096, 1985.
233. Grabowski, P. J., Seiter, S. R., and Sharp, P. A., A multicomponent complex is involved in the splicing of messenger RNA precursors, *Cell*, 42, 345, 1985.
234. Binderef, A. and Green, M. R., An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly, *EMBO J.*, 6, 2415, 1987.
235. Frendeway, D. and Keller, W., Stepwise assembly of a pre- mRNA splicing complex requires U snRNPs and specific intron sequences, *Cell*, 42, 355, 1985.
236. Keller, W., The RNA lariat: a new ring to the splicing of mRNA precursors, *Cell*, 39, 423, 1984.
237. Zeitlin, S. and Efstratiadis, A., In vivo splicing products of the rabbit B-globin pre mRNA, *Cell*, 39, 589, 1984.
238. Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E., and Abelson, J., Lariat structures are in vivo intermediates in yeast pre-mRNA splicing, *Cell*, 39, 611, 1984.
239. Wieringa, B., Hofer, E., and Weissmann, C., A minimal intron length but no specific internal sequence is required for splicing the large rabbit B-globin intron, *Cell*, 37, 915, 1984.
240. Rodriguez, J. R., Pikielny, C. W., and Rosbash, M., In vivo characterization of yeast mRNA processing intermediates, *Cell*, 39, 603, 1984.
241. Aebi, M., Hornig, H., Padgett, R. A., Reiser, J., and Weissmann, C., Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA, *Cell*, 47, 555, 1986.
242. Aebi, M., Hornig, H., and Weissmann, C., 5' Cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU, *Cell*, 50, 237, 1987.
243. Ruskin, B. and Green, M. R., Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing, *Nature*, 317, 732, 1985.
244. Reed, R. and Maniatis, T., Intron sequences involved in lariat formation during pre-mRNA splicing, *Cell*, 41, 95, 1985.
245. Jacquier, A., Rodriguez, J. R., and Rosbash, M., A quantitative analysis of the effects of 5' junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing, *Cell*, 43, 423, 1985.
246. Newman, A. J., Lin, R.-J., Cheng, S.-C., and Abelson, J., Molecular consequences of specific intron mutations on yeast mRNA splicing in vivo and in vitro, *Cell*, 42, 335, 1985.
247. Reed, R., Griffith, J., and Maniatis, T., Purification and visualization of native spliceosomes, *Cell*, 53, 949, 1988.
248. Treisman, R., Orkin, S. H., and Maniatis, T., Specific transcription and RNA splicing defects in five cloned beta-thalassemia genes, *Nature*, 302, 591, 1983.
249. Brody, E. and Abelson, J., The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction, *Science*, 228, 963, 1985.
250. Deimel, B., Louis, C., and Sekeris, C. E., The presence of small molecular weight RNAs in nuclear ribonucleoprotein particles containing hnRNA, *FEBS Lett.*, 73, 775, 1977.
251. Calvet, J. P. and Pederson, T., Base-pairing interactions between small nuclear RNAs and nuclear precursors as revealed by psoralen cross-linking in vivo, *Cell*, 26, 363, 1981.
252. Calvet, J. P., Mayer, L. M., and Pederson, T., Small nuclear RNA U2 is based paired to heterogeneous nuclear RNA, *Science*, 217, 456, 1982.
253. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., and Steitz, J. A., Are snRNPs involved in splicing?, *Nature*, 283, 220, 1980.
254. Rogers, J. and Wall, R., A mechanism for RNA splicing, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 1877, 1980.
255. Fradin, A., Jove, R., Hemenway, C., Keiser, H. D., Manley, M. L., and Prives, C., Splicing pathways of SV40 mRNAs in *X. laevis* oocytes differ in their requirements for snRNPs, *Cell*, 37, 927, 1984.
256. Grabowski, P. J. and Sharp, P. A., Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small ribonucleoprotein particles in the spliceosome, *Science*, 233, 1294-1299, 1986.
257. Pikielny, C. W., Rymond, B. C., and Rosbash, M., Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes, *Nature*, 324, 341, 1986.
258. Cheng, S. C. and Abelson, J., Spliceosome assembly in yeast, *Genes Dev.*, 1, 1014, 1987.
258a. Seraphin, B. and Rosbash, M., Identification of functional U1 snRNA-Pre-mRNA complexes committed to spliceosome assembly and splicing, *Cell*, 59, 349, 1989.
259. Zillman, M., Rose, S. A., and Berget, S. M., U1 small nuclear ribonucleoproteins are required early during spliceosome assembly, *Mol. Cell. Biol.*, 7, 2877, 1987.
260. Zillman, M., Zapp, M. L., and Berget, S. M., Gel electrophoresis of splicing complexes containing U1 small nuclear ribonucleoprotein particles, *Mol. Cell. Biol.*, 8, 814, 1988.
261. Konarska, M. M. and Sharp, P. A., Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs, *Cell*, 46, 845, 1986.
262. Kramer, A., Analysis of RNase-A-resistant regions of Adenovirus 2 major late precursor-mRNA in splicing extracts reveals an ordered interaction of nuclear components with the substrate RNA, *J. Mol. Biol.*, 196, 559, 1987.
263. Chabot, B., Black, D. L., LeMaire, D. M., and Steitz, J. A., The 3' splice site in pre-messenger RNA is recognized by a small nuclear ribonucleoprotein, *Science*, 230, 1344, 1985.
264. Ruskin, B., Zamore, P. D., and Green, M. R., A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly, *Cell*, 52, 207, 1988.
265. Berget, S. M. and Robberson, B. L., U1, U2, and U4/U6 small nuclear ribonucleoproteins are required for in vitro splicing but not polyadenylation, *Cell*, 46, 691, 1986.
266. Black, D. L. and Steitz, J. A., Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoproteins, *Cell*, 46, 697, 1986.
267. Black, D. L., Chabot, B., and Steitz, J. A., U2 as well as U1 small nuclear ribonucleoproteins are involved in pre-messenger RNA splicing, *Cell*, 42, 737, 1985.
268. Konarska, M. M. and Sharp, P. A., Association of U2, U4, U5, and U6 small nuclear ribonucleoproteins in a spliceosome-type complex in absence of precursor RNA, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 5459, 1988.
269. Binderef, A. and Green, M. R., Ribonucleoprotein complex formation during pre-mRNA splicing in vitro, *Mol. Cell. Biol.*, 6, 2382, 1986.
269a. Blencowe, B. J., Sproat, B. S., Ryder, U., Barabino, S., and Lamond, A. I., Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides, *Cell*, 59, 531, 1989.
270. Langford, C. J., Klinz, F.-J., Donath, C., and Gallwitz, D., Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast, *Cell*, 36, 645, 1984.
271. Seiyono, B. and Pederson, T., Ribonucleoprotein Organization of RNA. XXX. Evidence that U1 small nuclear RNA is a ribonucleoprotein when base-paired with pre-messenger RNA in vivo, *J. Mol. Biol.*, 174, 285, 1984.
272. Padgett, R. A., Mount, S. A., Steitz, J. A., and Sharp, P. A., Splicing of messenger RNA precursors is inhibited by antisera to small nuclear ribonucleoproteins, *Cell*, 35, 101, 1983.
273. Tatei, K., Takemura, K., Mayeda, A., Fujiwara, Y., Tanaka, H., Ishiharna, A., and Ohshima, K., U1 RNA-protein complex preferentially binds to both 5' and 3' splice junction sequences in RNA or single stranded DNA, *Proc. Natl Acad. Sci. U.S.A.*, 81, 6281, 1984.
274. Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J. A., The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro, *Cell*, 33, 509, 1983.
275. Kramer, A., Keller, W., Appel, B., and Luhrmann, R., The 5' terminus

of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors, *Cell*, 38, 299, 1984.

276. Krainer, A. R. and Maniatis, T., Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-RNA splicing in vitro, *Cell*, 42, 725, 1985.

276a. Zhuang, Y. and Weiner, A. M., A compensatory base change in human U2 snRNA can suppress a branch site mutation, *Genes Dev.*, 3, 1545, 1989.

276b. Hamm, J., Darzynkiewicz, E., Tahara, S. M., and Mattaj, I. W., Functional analysis of mutant *Xenopus* U2 snRNAs, *Cell*, 55, 159, 1988.

277. Zhuang, Y. and Weiner, A. M., A compensatory base change in U1 snRNA suppresses a 5' splice site mutation, *Cell*, 46, 827, 1986.

278. Parker, R., Siliciano, P. G., and Guthrie, C., Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA, *Cell*, 49, 229, 1987.

279. Ruskin, B. and Green, M. R., Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing, *Cell*, 43, 131, 1985.

280. Keller, E. B. and Noon, W. A., Intron splicing: a conserved internal signal in introns of animal pre-mRNAs, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 7417, 1984.

281. Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J., and Guthrie, C., Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome, *EMBO J.*, 5, 1683, 1986.

282. Birnstiel, M. L., Busslinger, M., and Strub, K., Transcription termination and 3' processing: the end is in site!, *Cell*, 41, 349, 1985.

283. Galli, G., Hofstetter, H., Stunnenberg, H. G., and Birnstiel, M. L., Biochemical complementation with RNA in the *Xenopus* oocyte: a small RNA is required for the generation of 3' histone mRNA termini, *Cell*, 34, 823, 1983.

284. Gilmartin, G. M., Schaufele, F., Schaffner, G., and Birnstiel, M. L., Functional analysis of the sea urchin U7 small nuclear RNA., *Mol. Cell. Biol.*, 8, 1075, 1988.

285. Schaufele, F., Gilmartin, G. M., Banerjee, W., and Birnstiel, M. L., Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of messenger RNA, *Nature (London)*, 323, 777, 1986.

286. Mowry, K. L. and Steitz, J. A., Both conserved signals on mammalian histone pre-mRNAs associate with small nuclear ribonucleoproteins during 3' end formation in vitro, *Mol. Cell. Biol.*, 7, 1663, 1987.

287. Cotten, M., Gick, O., Vasserot, A., Schaffner, G., and Birnstiel, M. L., Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the in vitro 3' RNA processing reaction., *EMBO J.*, 7, 801, 1988.

288. Moore, C. L. and Sharp, P. A., Accurate cleavage and polyadenylation of exogenous RNA substrate, *Cell*, 41, 845, 1985.

289. Stefano, J. E. and Adams, D. E., Assembly of a polyadenylation-specific 25S ribonucleoprotein complex in vitro, *Mol. Cell. Biol.*, 8, 2052, 1988.

290. Moore, C. L., Chen, J., and Whoriskey, J., Two proteins cross-linked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding, *EMBO J.*, 7, 3159, 1988.

291. Berget, S. M., Are U4 small nuclear ribonucleoproteins involved in polyadenylation?, *Nature*, 309, 179, 1984.

292. Tague, B. W. and Gerbi, S. A., Processing of the large rRNA precursor: two proposed categories of RNA-RNA interaction in eukaryotes, *J. Mol. Evol.*, 20, 362, 1984.